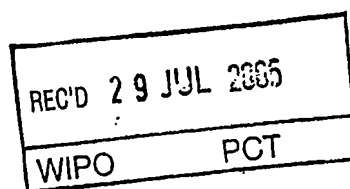


IB 05/1647

日本国特許庁  
JAPAN PATENT OFFICE



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application: 2004年 5月11日

出願番号  
Application Number: 特願2004-232811

パリ条約による外国への出願  
に用いる優先権の主張の基礎  
となる出願の国コードと出願  
番号

The country code and number  
of your priority application,  
to be used for filing abroad  
under the Paris Convention, is

J P 2004-232811

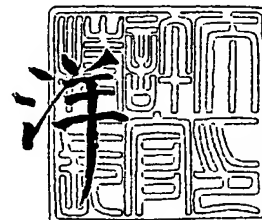
出願人  
Applicant(s): 株式会社RNAi

PRIORITY DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

2005年 6月30日

特許庁長官  
Commissioner,  
Japan Patent Office

小川



本証明書には以下のCD-Rが付属している。

ファイル名 : 2004232811seq1.txt

ファイルサイズ : 177MB

記録日 : 2005年7月7日

【書類名】 特許願  
【整理番号】 PNYA-16219  
【提出日】 平成16年 5月11日  
【あて先】 特許庁長官殿  
【国際特許分類】 C12N 15/00  
【発明者】  
    【住所又は居所】 東京都千代田区一番町 6-4-102  
    【氏名】 内藤 雄樹  
【発明者】  
    【住所又は居所】 東京都文京区西片二丁目 25-8 モンテベルデ本郷西片 303号  
    【氏名】 藤野 真人  
【発明者】  
    【住所又は居所】 東京都中野区若宮 1丁目 28番 3号  
    【氏名】 小口 しのぶ  
【発明者】  
    【住所又は居所】 神奈川県横浜市西区みなとみらい四丁目 10番 1-E 1706号  
    【氏名】 名取 幸和  
【特許出願人】  
    【住所又は居所】 東京都文京区本郷三丁目 32番地 6-703号  
    【氏名又は名称】 株式会社RNAi  
【代理人】  
    【識別番号】 100089118  
    【弁理士】  
    【氏名又は名称】 酒井 宏明  
【手数料の表示】  
    【予納台帳番号】 036711  
    【納付金額】 16,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1

特願 2004-232811

ページ: 1/E

【書類名】 特許願  
【整理番号】 PNYA-16219  
【提出日】 平成16年 5月11日  
【あて先】 特許庁長官殿  
【国際特許分類】 C12N 15/00  
【発明者】  
    【住所又は居所】 東京都千代田区一番町 6-4-102  
    【氏名】 内藤 雄樹  
【発明者】  
    【住所又は居所】 東京都文京区西片二丁目 25-8 モンテベルデ本郷西片 303  
    【氏名】 藤野 真人  
【発明者】  
    【住所又は居所】 東京都中野区若宮 1丁目 28番 3号  
    【氏名】 小口 しのぶ  
【発明者】  
    【住所又は居所】 神奈川県横浜市西区みなとみらい四丁目 10番 1-E1706号  
    【氏名】 名取 幸和  
【特許出願人】  
    【住所又は居所】 東京都文京区本郷三丁目 32番地 6-703号  
    【氏名又は名称】 株式会社RNA i  
【代理人】  
    【識別番号】 100089118  
    【弁理士】  
    【氏名又は名称】 酒井 宏明  
【手数料の表示】  
    【予納台帳番号】 036711  
    【納付金額】 16,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1

【書類名】 特許願  
【整理番号】 PNYA-16219  
【提出日】 平成16年 5月11日  
【あて先】 特許庁長官殿  
【国際特許分類】 C12N 15/00  
【発明者】  
    【住所又は居所】 東京都千代田区一番町 6-4-102  
    【氏名】 内藤 雄樹  
【発明者】  
    【住所又は居所】 東京都文京区西片二丁目 25-8 モンテベルデ本郷西片 303号  
    【氏名】 藤野 真人  
【発明者】  
    【住所又は居所】 東京都中野区若宮 1丁目 28番3号  
    【氏名】 小口 しのぶ  
【発明者】  
    【住所又は居所】 神奈川県横浜市西区みなとみらい四丁目 10番 1-E 1706号  
    【氏名】 名取 幸和  
【特許出願人】  
    【住所又は居所】 東京都文京区本郷三丁目 32番地 6-703号  
    【氏名又は名称】 株式会社RNAi  
【代理人】  
    【識別番号】 100089118  
    【弁理士】  
    【氏名又は名称】 酒井 宏明  
【手数料の表示】  
    【予納台帳番号】 036711  
    【納付金額】 16,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1

## 【書類名】特許請求の範囲

## 【請求項 1】

目的生物の遺伝子の中から選ばれる標的遺伝子について、RNA干渉を生じさせるポリヌクレオチドであって、

少なくとも 2 本鎖領域を有し、

前記 2 本鎖領域における一方の鎖が、標的遺伝子の塩基配列中に含まれる、下記 (a) から (d) の規則に従う規定配列と相同な塩基配列からなり、

前記 2 本鎖領域における他方の鎖が、前記規定配列と相同な塩基配列と相補的な配列を有する塩基配列からなることを特徴とするポリヌクレオチド。

(a) 3'末端の塩基が、アデニン、チミンまたはウラシルである。

(b) 5'末端の塩基が、グアニンまたはシトシンである。

(c) 3'末端の 7 塩基の配列において、アデニン、チミンおよびウラシルからなる群より選ばれる 1 種または 2 種以上の塩基がリッチである。

(d) 塩基数が、細胞毒性を生じさせずに RNA 干渉を生じさせ得る数である。

## 【請求項 2】

前記規定配列と相同な塩基配列の少なくとも 80% 以上の塩基が、前記規定配列の塩基配列と一致することを特徴とする請求項 1 に記載のポリヌクレオチド。

## 【請求項 3】

前記規則 (c) において、7 塩基のうち少なくとも 3 塩基以上がアデニン、チミンおよびウラシルからなる群より選ばれる 1 種または 2 種以上であることを特徴とする請求項 1 または 2 に記載のポリヌクレオチド。

## 【請求項 4】

前記規則 (d) において、塩基数が 13 ~ 28 であることを特徴とする請求項 1 ~ 3 のいずれか一項に記載のポリヌクレオチド。

## 【請求項 5】

前記規定配列が、さらに下記 (e) の規則に従う配列であることを特徴とする請求項 1 ~ 4 のいずれか一項に記載のポリヌクレオチド。

(e) 10 塩基以上グアニンまたはシトシンが連続する配列を含まない。

## 【請求項 6】

前記規定配列が、さらに下記 (f) の規則に従う配列であることを特徴とする請求項 1 ~ 5 のいずれか一項に記載のポリヌクレオチド。

(f) 目的生物の全遺伝子配列のうち、標的遺伝子以外の他の遺伝子の塩基配列中に、当該規定配列と 90% 以上の相同性を有する配列が含まれない。

## 【請求項 7】

前記規定配列が、配列表配列番号 47 ~ 配列番号 817081 のいずれかに記載の塩基配列からなることを特徴とする請求項 1 ~ 6 のいずれか一項に記載のポリヌクレオチド。

## 【請求項 8】

2 本鎖ポリヌクレオチドであることを特徴とする請求項 1 ~ 7 のいずれか一項に記載のポリヌクレオチド。

## 【請求項 9】

前記 2 本鎖ポリヌクレオチドの一方の鎖が、前記規定配列と相同な塩基配列の 3'末端にオーバーハング部を有する塩基配列からなり、

前記 2 本鎖ポリヌクレオチドの他方の鎖が、前記規定配列と相同な塩基配列と相補的な配列の 3'末端にオーバーハング部を有する塩基配列からなることを特徴とする請求項 8 に記載のポリヌクレオチド。

## 【請求項 10】

ヘアピン構造を有する 1 本鎖ポリヌクレオチドであり、

前記 2 本鎖領域における一方の鎖の 3'末端と、前記 2 本鎖領域における他方の鎖の 5'末端とを連結するループ部を有することを特徴とする請求項 1 ~ 7 のいずれか一項に記載のポリヌクレオチド。

【請求項 11】

請求項 1 ～ 10 のいずれか一項に記載のポリヌクレオチドを、発現を抑制しようとする標的遺伝子の発現系に導入して標的遺伝子の発現を抑制することを特徴とする遺伝子発現抑制方法。

## 【書類名】明細書

【発明の名称】RNA干渉を生じさせるポリヌクレオチド、および、これを用いた遺伝子発現抑制方法

## 【技術分野】

## 【0001】

本発明は、RNA干渉を生じさせるポリヌクレオチドに関する。以下本明細書においてRNA干渉のことを「RNAi」と表記する場合がある。

## 【背景技術】

## 【0002】

RNA干渉は、機能阻害したい遺伝子の特定領域と相同なセンスRNAとアンチセンスRNAからなる2本鎖RNA (double-stranded RNA、以下「dsRNA」という) が標的遺伝子の転写産物であるmRNAの相同部分を干渉破壊するという現象で、1998年に線虫を用いた実験により初めて提唱された。しかし、哺乳類においては、約30塩基対以上の長いdsRNAを細胞内へ導入すると、インターフェロンレスポンスが誘導され、細胞がアポトーシスによって死んでしまうため、RNAi法を用いることが困難であった。

## 【0003】

一方、マウス初期胚や哺乳類培養細胞では、RNA干渉が起こりえることが示され、RNA干渉の誘導機構そのものは、哺乳類細胞にも存在することがわかってきた。現在では、およそ21~23塩基対の短い2本鎖RNA (short interfering RNA、siRNA) が、哺乳類細胞系でも細胞毒性を示さずにRNA干渉を誘導できることが示され、哺乳類においてもRNAi法を利用することが可能となってきた。

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0004】

RNAi法は様々な応用が期待される技術である。しかし、ショウジョウバエや線虫では、ある遺伝子の特定領域と相同なdsRNAおよびsiRNAは、ほとんどの配列でRNA干渉効果を示すのに対して、哺乳類では無作為に選択した(21塩基の)siRNAの70~80%はRNA干渉効果を示さない。これは、哺乳類においてRNAi法を用いた遺伝子機能解析を行う際に大きな問題点となっている。

## 【0005】

また、従来siRNAを設計するにあたっては、試験者等の経験やセンスに依存する部分が大きく、実際にRNA干渉効果を示すsiRNAを高い確率で設計することが困難であった。さらに、RNA合成は、費用、時間等を要することから、RNA干渉を行うために無駄なsiRNAを合成してしまうことは、RNA干渉のさらなる研究や、RNA干渉を利用した数々の利用法の普及を妨げる要因となっていた。

## 【課題を解決するための手段】

## 【0006】

本発明者らは、上記課題を解決するために、RNAi法を用いる際最も労力、時間、費用を要する部分の1つである、siRNAの入手を容易に行う手法について検討を進めた。siRNAの調製は哺乳類において特に問題となっていることから、本発明者らは哺乳類培養細胞系を用いて、RNA干渉に有効なsiRNAの配列規則性を同定することを試みたところ、有効なsiRNAの配列には所定の規則性があることを見だし、本発明を完成するに至った。すなわち、本発明は、次の通りである。

## 【0007】

〔1〕 目的生物の遺伝子の中から選ばれる標的遺伝子について、RNA干渉を生じさせるポリヌクレオチドであって、

少なくとも2本鎖領域を有し、

前記2本鎖領域における一方の鎖が、標的遺伝子の塩基配列中に含まれる、下記(a)から(d)の規則に従う規定配列と相同な塩基配列からなり、



前記2本鎖領域における他方の鎖が、前記規定配列と相同な塩基配列と相補的な配列を有する塩基配列からなることを特徴とするポリヌクレオチド。

(a) 3'末端の塩基が、アデニン、チミンまたはウラシルである。

(b) 5'末端の塩基が、グアニンまたはシトシンである。

(c) 3'末端の7塩基の配列において、アデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上の塩基がリッチである。

(d) 塩基数が、細胞毒性を生じさせずにRNA干渉を生じさせ得る数である。

【0008】

[2] 前記規定配列と相同な塩基配列の少なくとも80%以上の塩基が、前記規定配列の塩基配列と一致することを特徴とする[1]に記載のポリヌクレオチド。

【0009】

[3] 前記規則(c)において、7塩基のうち少なくとも3塩基以上がアデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上であることを特徴とする[1]または[2]に記載のポリヌクレオチド。

【0010】

[4] 前記規則(d)において、塩基数が13~28であることを特徴とする[1]~[3]のいずれか一項に記載のポリヌクレオチド。

【0011】

[5] 前記規定配列が、さらに下記(e)の規則に従う配列であることを特徴とする[1]~[4]のいずれか一項に記載のポリヌクレオチド。

(e) 10塩基以上グアニンまたはシトシンが連続する配列を含まない。

【0012】

[6] 前記規定配列が、さらに下記(f)の規則に従う配列であることを特徴とする[1]~[5]のいずれか一項に記載のポリヌクレオチド。

(f) 目的生物の全遺伝子配列のうち、標的遺伝子以外の他の遺伝子の塩基配列中に、当該規定配列と90%以上の相同性を有する配列が含まれない。

【0013】

[7] 前記規定配列が、配列表配列番号47~配列番号817081のいずれかに記載の塩基配列からなることを特徴とする[1]~[6]のいずれか一項に記載のポリヌクレオチド。

【0014】

[8] 2本鎖ポリヌクレオチドであることを特徴とする[1]~[7]のいずれか一項に記載のポリヌクレオチド。

【0015】

[9] 前記2本鎖ポリヌクレオチドの一方の鎖が、前記規定配列と相同な塩基配列の3'末端にオーバーハング部を有する塩基配列からなり、

前記2本鎖ポリヌクレオチドの他方の鎖が、前記規定配列と相同な塩基配列と相補的な配列の3'末端にオーバーハング部を有する塩基配列からなることを特徴とする[8]に記載のポリヌクレオチド。

【0016】

[10] ヘアピン構造を有する1本鎖ポリヌクレオチドであり、

前記2本鎖領域における一方の鎖の3'末端と、前記2本鎖領域における他方の鎖の5'末端とを連結するループ部を有することを特徴とする[1]~[7]のいずれか一項に記載のポリヌクレオチド。

【0017】

[11] [1]~[10]のいずれか一項に記載のポリヌクレオチドを、発現を抑制しようとする標的遺伝子の発現系に導入して標的遺伝子の発現を抑制することを特徴とする遺伝子発現抑制方法。

【発明の効果】

【0018】

本発明のポリヌクレオチドは、標的遺伝子に対して高いRNA干渉効果を有するとともに、標的遺伝子と関係のない遺伝子についてRNA干渉を生じる危険性が非常に小さいため、発現を抑制しようとする標的遺伝子だけに特異的にRNA干渉を生じさせることができる。したがって、本発明は、RNA干渉を利用する試験、治療方法等に好適に利用することができ、哺乳類などの高等動物、特にヒトを対象とするRNA干渉を行う際に特に有効性を発揮するものである。

【発明を実施するための最良の形態】

【0019】

本発明の実施の形態を、以下の<1>から<7>の順に従って説明する。

- <1> RNA干渉の標的塩基配列検索方法
- <2> RNA干渉を生じさせるポリヌクレオチドの塩基配列設計方法
- <3> ポリヌクレオチドの作製方法
- <4> 遺伝子の発現抑制方法
- <5> siRNA配列設計プログラム
- <6> siRNA配列設計ビジネスモデルシステム
- <7> siRNA配列設計プログラムを実行する塩基配列処理装置等

【0020】

- <1> RNA干渉の標的塩基配列検索方法

本発明の検索方法は、目的生物の遺伝子の中から選ばれる標的遺伝子の塩基配列中から、RNA干渉の起因となる塩基配列を探し出す方法である。RNA干渉を生じさせる対象となる目的生物は、特に限定されず、大腸菌をはじめとする原核生物、酵母、真菌等の微生物、哺乳類をはじめとする動物、昆虫、植物等のいずれでも良い。

【0021】

具体的には、本発明の検索方法では、RNA干渉の標的遺伝子の塩基配列から、下記（a）から（d）の規則に従う配列部位を検索する。

- （a） 3'末端の塩基が、アデニン、チミンまたはウラシルである。
- （b） 5'末端の塩基が、グアニンまたはシトシンである。
- （c） 3'末端の7塩基の配列において、アデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上の塩基がリッチである。
- （d） 塩基数が、細胞毒性を生じさせずにRNA干渉を生じさせ得る数である。

【0022】

「標的遺伝子」における「遺伝子」とは、遺伝情報をコードする媒体のことという。「遺伝子」はDNA、RNA、DNAおよびRNAの複合体など遺伝情報をコードする物質により構成されるが、遺伝情報としては物質そのものではなく塩基配列を電子データ化したものをコンピュータ上などで扱うことが可能である。「標的遺伝子」は、1つのコード領域としてもよいし、複数のコード領域に渡り標的としてもよいし、また、配列が判明したすべてのポリヌクレオチドを標的としてもよい。特定の機能を有する遺伝子について検索したい場合には、その特定遺伝子のみを標的とすることにより、当該特定遺伝子に特異的にRNA干渉を生じさせる塩基配列を効率よく検索することができる。すなわち、RNA干渉はmRNAを干渉破壊する現象として知られており、特定のコード領域を選択することにより検索負荷を軽減できる。また、転写領域のひとまとまりを標的遺伝子として検索してもよい。なお、本明細書において塩基配列は特に断らない限りセンス鎖、すなわち、mRNAの配列を基準として示す。また、本明細書中では上記（a）から（d）の規則を満たす塩基配列のことを「規定配列」という。上記規則においては、塩基配列がDNAの配列であればチミンが、RNAの配列であればウラシルが対応する。

【0023】

規則（c）は、3'末端近傍の配列にアデニン、チミン、およびウラシルからなる群より選ばれる塩基がリッチに含まれていることを規定しており、具体的に検索を行う際の一指標として3'末端部から7塩基の範囲内において、アデニン、チミン、およびウラシルから選ばれる塩基がリッチな配列であることを規定している。

## 【0024】

規則(c)において、「リッチな配列」とは特定の塩基が現れる頻度が高いことを意味し、割合を概略的に示すと、規定配列における3'末端側の5～10塩基、好ましくは7塩基の配列中にアデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上が、少なくとも40%以上、より好ましくは50%以上含まれることを意味する。より具体的には、例えば約19塩基程度の規定配列の場合を例に挙げると、3'末端側の7塩基のうち好ましくは少なくとも3塩基以上、より好ましくは4塩基以上、特に好ましくは5塩基以上が、アデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上である。

## 【0025】

規則(c)に該当するか否かを確認する手段は特に制限はなく、7塩基中の好ましくは3塩基以上、より好ましくは4塩基以上、特に好ましくは5塩基以上がアデニン、チミンまたはウラシルであることを確認できればよい。例えば、アデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上が、3'末端側の7塩基の配列中に3つ以上含まれることをリッチであると定義した場合を例として説明すると、3'末端の第1番目の塩基から逐次上記3種の塩基のいずれかであるか照合し、第7番目の塩基に至るまでに3つ現れた場合に、規則(c)に適合すると判断することができる。例えば、第3番目までに3つ現れれば、3つの塩基を調べれば足りる。すなわち、規則(c)について検索する場合、必ずしも3'末端側の7塩基についてすべて照合することは要しない。逆に第7番目までに3つ以上現れなければリッチではなく、規則(c)を満たさないと判断される。

## 【0026】

二本鎖ポリヌクレオチドにおいてはアデニンとチミンまたはウラシルとが相補的に水素結合することは周知である。また、グアニンとシトシンとの相補的な水素結合(G-C水素結合)においては3つの水素結合部位が形成されるのに対し、アデニンとチミンまたはウラシルとの相補的水素結合(A-(T/U)水素結合)においては2つの水素結合部位からなり、一般的に言ってG-C水素結合に対し、A-(T/U)水素結合のほうが結合力は弱い。

## 【0027】

規則(d)においては、検索する塩基配列の塩基数を規定している。検索する塩基配列の塩基数は、RNA干渉を生じさせ得る塩基数である。また、生物の種類などの条件により、塩基数があまりに大きすぎるsiRNAでは細胞毒性を生じてしまうことが知られている。塩基数の上限は、RNA干渉を生じさせようとする生物の種類などにより異なるが、siRNAを構成する一本鎖の塩基数はいずれの種にせよ30以下であることが好ましい。また、哺乳動物の塩基数にあつては、好ましくは24以下、より好ましくは22以下である。また、下限はRNA干渉を生じさせる限りにおいて特に制限されるものではないが、好ましくは15以上、より好ましくは18以上、さらに好ましくは20以上である。siRNAを構成する一本鎖としての塩基数は、21で検索することが特に好ましい。

## 【0028】

なお、下記にても説明するが、siRNAには、規定配列の3'末端にオーバーハング部が設けられる。オーバーハング部は塩基数2であることが好適である。したがって、オーバーハング部を含めず、規定配列のみの塩基数の上限としては、好ましくは28以下、より好ましくは22以下、さらに好ましくは20以下であり、下限としては、好ましくは13以上、より好ましくは16以上、さらに好ましくは18以上である。規定配列の最も好ましい塩基数は19である。RNAiの標的塩基配列の検索は、オーバーハング部を含める場合および含めない場合のいずれで検索してもよい。

## 【0029】

規定配列に従う塩基配列は、RNA干渉を生じさせる確率が極めて高い。したがって、本発明の検索方法により、RNA干渉を生じさせる配列を極めて高い確率で検索することが可能であり、RNA干渉を生じさせるポリヌクレオチドの設計を簡便化することができ

る。

【0030】

また、他の好ましい例として、規定配列が、さらに下記の規則(e)に従う配列であることが挙げられる。

(e) 10塩基以上グアニンまたはシトシンが連続する配列を含まない。

【0031】

規則(e)においては、検索する塩基配列が、10塩基以上グアニン(G)および/またはシトシン(C)が連続した配列を含まない配列であることを規定している。10塩基以上グアニンおよび/またはシトシンが連続するというのは、例えば、グアニンまたはシトシンの一方のみが連続する場合と、グアニンおよびシトシンとが混在する配列となっている場合の双方を含み、より具体的には、GGGGGGGGGG、CCCCCCCCCのほか、GおよびCの混合配列であるGCGCCCGCGなども含まれる。

【0032】

また、標的遺伝子と関係のない遺伝子についてまでRNA干渉を生じさせないようにするためには、設計した配列と同一または類似の配列が他の遺伝子に含まれていないか検索することが好ましい。設計した配列と同一または類似の配列の検索は、一般的なホモロジー検索を行うことができるソフトウェア等を用いて行えばよい。この際、siRNAの2本の鎖(センス鎖およびアンチセンス鎖)によるRNAi効果を考慮して、「設計した配列」および「設計した配列と相補的な塩基配列を有する配列(相補配列)」の両方について、同一または類似の配列が他の遺伝子に含まれていないか検索することがより好ましい。設計した配列のうち、設計した配列またはその相補配列と同一/類似配列を有する配列を除外することにより、標的とする遺伝子のみに特異的にRNA干渉を生じさせる配列を設計することができる。

【0033】

したがって、設計した配列のうち、他の遺伝子に塩基配列のミスマッチ数が小さい類似配列が存在するような配列を除外することにより、特異性の高い配列を選別することができる。例えば、塩基数19の塩基配列を設計する場合、他の遺伝子にミスマッチが2塩基以下の類似配列が存在する配列は、除外することが好ましい。ここで類似性判断の閾値となるミスマッチ数の値を大きく設定するほど、設計する配列の特異性が高まる。塩基数19の塩基配列を設計する場合、より好ましくは、他の遺伝子にミスマッチが3塩基以下の類似配列が存在する配列を除外することが好ましく、さらに好ましくは、他の遺伝子にミスマッチが4塩基以下の類似配列が存在する配列を除外することが好ましい。また、当該規定配列を有する配列およびその相補配列の両方について、塩基配列のミスマッチ数が小さい類似配列が他の遺伝子に存在するような配列を除外することにより、より特異性の高い配列を設計することができるため、好ましい。

【0034】

配列の類似性を判断する基準となるミスマッチ数は、設計する配列の塩基数によっても異なるため、一概に規定することは難しい。塩基配列のミスマッチ数を相同性で規定することとすると、下記の規則(f)に従う配列であるか検索すればよい。

(f) 目的生物の全遺伝子配列のうち、標的遺伝子以外の他の遺伝子の塩基配列中に、当該規定配列を有する配列と90%以上の相同性を有する配列が含まれない。

【0035】

規則(f)において、標的遺伝子以外の他の遺伝子の塩基配列中に、規定配列と85%以上の相同性を有する配列が含まれないことが好ましく、規定配列と80%以上の相同性を有する配列が含まれないことがより好ましく、規定配列と75%以上の相同性を有する配列が含まれないことが好ましい。また、ここで、当該規定配列を有する配列およびその相補配列の両方について、塩基配列の相同性が高い類似配列が他の遺伝子に存在するような配列を除外することにより、より特異性の高い配列を設計することができるため、好ましい。

【0036】

なお、規定配列の検索は、塩基数を定めた上で上記の(a)から(c)の規則などに従う部分を検索するようなプログラムを搭載するコンピュータを用いて効率的に検出可能である。より具体的な実施の形態は下記<5> siRNA配列設計プログラム、および、<7> siRNA配列設計プログラムを実行する塩基配列処理装置等の欄に示す。

#### 【0037】

<2> RNA干渉を生じさせるポリヌクレオチドの塩基配列設計方法

本発明の塩基配列設計方法は、上記の検索方法により検索された塩基配列に基づいてRNA干渉を生じさせるポリヌクレオチドの塩基配列を設計するものである。RNA干渉を生じさせるポリヌクレオチドは、上記の検索方法により検索された規定配列に基づいて設計された2本鎖領域を有するポリヌクレオチドであり、標的遺伝子について、RNA干渉を生じさせることができるものであれば、特に限定されない。

#### 【0038】

RNA干渉を生じさせるポリヌクレオチドとしては、主として、siRNA等の2本鎖ポリヌクレオチド型と、ヘアピン構造を有するRNA(short hairpin RNA: shRNA)等の1本鎖ポリヌクレオチド型に分類することができる。

#### 【0039】

siRNAやshRNAは、主としてRNAからなるが、一部にDNAが含まれている混成ポリヌクレオチドも含まれる。本発明の塩基配列設計方法では、上記(a)から(d)の規則に従う塩基配列を標的遺伝子の塩基配列から検索し、検索された塩基配列と相同な塩基配列を設計する。また、他の好ましい設計例としては、上記(e)および(f)の規則などを考慮してもよい。(a)から(d)の規則および検索の手法などについては、上記本発明の検索方法について説明したとおりである。

#### 【0040】

RNA干渉を生じさせるポリヌクレオチドにおける2本鎖領域は、一方の鎖が、標的遺伝子の塩基配列中に含まれる、上記(a)から(d)の規則に従う規定配列と相同な塩基配列からなり、他方の鎖が、前記規定配列と相同な塩基配列と相補的な配列を有する塩基配列からなる。「相同な配列」とは、同一の配列および当該RNA干渉を生じさせるという機能を失わない範囲で同一配列に対し欠失、置換、挿入などの変異を含む配列のことをいう。標的遺伝子の種類、配列などの条件にもよるが、許容される変異を相同性(ホモロジー)で例示すると、好ましくは80%以上、より好ましくは90%以上、さらに好ましくは95%以上である。許容される変異の程度としての相同性を算出する場合、同一の検索アルゴリズムを用いて算出された数値どうしを比較することが望ましい。検索アルゴリズムは特に限定されないが、局所的な配列の検索に適したものが好適であり、より具体的にはBLAST、searchなどを好適に用いることができる。

#### 【0041】

上記のように、検索された配列は若干の改変が許容されるが、設計される塩基配列の塩基数は、検索された配列と同一とすることが特に好ましい。塩基数を同一とした場合について改変の許容度を例示すると、設計される塩基配列の塩基が、好ましくは80%以上、より好ましくは90%以上、特に好ましくは95%以上検索された配列と一致することが好適である。例えば、塩基数19の塩基配列を設計する場合であれば、好ましくは16塩基以上、より好ましくは18塩基以上が、検索された塩基配列と一致することが好ましい。また、検索された塩基配列と相同な配列を設計する場合、検索された塩基配列の3'末端の塩基と設計される塩基配列の3'末端の塩基とは同一であることが望ましく、また、検索された塩基配列の5'末端の塩基と設計される塩基配列の5'末端の塩基とは同一であることが望ましい。

#### 【0042】

通常siRNA分子には、オーバーハング部が設けられる。オーバーハング部とは、2本鎖のsiRNA分子において、各鎖の3'末端に設けられた、一本鎖の状態で突出した部分である。オーバーハング部は、生物の種類などにもよるが、塩基数は2が特に好適である。オーバーハング部の塩基配列は、基本的には任意であるが、検索元の標的遺伝子と

同一の塩基配列、TT、あるいはUUなどが好適に用いられる場合がある。上記のように検索された塩基配列と相同な配列となるように設計された規定配列の3'末端に、オーバーハング部を設けることにより、siRNAを構成するセンス鎖が設計される。

#### 【0043】

また、最初から規定配列およびオーバーハング部を含めて検索を行い、設計することもできる。オーバーハング部の好ましい塩基数は2である。したがって、例えば、塩基数19の規定配列および塩基数2のオーバーハング部からなるsiRNAを構成する一本鎖の設計をする場合には、オーバーハング部を含めたsiRNAの塩基数としては塩基数21の配列を標的遺伝子から検索すればよく、また2本鎖の状態について検索する場合には、塩基数23の配列を検索してもよい。

#### 【0044】

shRNAは、前記2本鎖領域における一方の鎖の3'末端と、前記2本鎖領域における他方の鎖の5'末端とをループ部で連結することによって1本鎖ポリヌクレオチドとしたものである。shRNAは、前記一方の鎖の5'末端および/または前記他方の鎖の3'末端に、一本鎖の状態で突出した部分を有していても良い。このようなshRNAは、WO01/49844等、公知の手法に従って設計することができる。

#### 【0045】

本発明の塩基配列設計方法では、上記のように所望の標的遺伝子から所定の配列を検索してくるが、RNA干渉を生じさせようとする対象は、標的遺伝子の由来と必ずしも一致せずともよく、類縁種などに適用可能である。例えば、第1の種から単離された遺伝子を標的遺伝子とし、第1の種の類縁種である第2の種に用いるsiRNAを設計することもできる。さらに、例えば複数種の哺乳類から共通配列を検索し、この共通配列から上記規定配列を検索して設計することにより、哺乳類に幅広く適用可能なsiRNAの設計が可能である。複数の哺乳類に共通する配列は、他の哺乳類においても保存されている確率が高いと考えられるためである。

#### 【0046】

本発明の設計方法により、RNA干渉を生じさせるRNA分子を、高い確率でしかも容易に設計することができる。RNAの合成は、未だ労力、時間、費用を要するが、本発明の設計方法によりそれらを大幅に軽減することが可能である。

#### 【0047】

##### <3>ポリヌクレオチドの作製方法

本発明のポリヌクレオチドの作製方法は、RNA干渉を生じさせる確率の高いポリヌクレオチドを作製する方法である。本発明のポリヌクレオチドは、ポリヌクレオチドの塩基配列が上記本発明の塩基配列設計方法に従って設計され、その配列設計に従うようにポリヌクレオチドが合成される。本発明のポリヌクレオチドには、上述のように、siRNA等の2本鎖ポリヌクレオチド型と、shRNA等の1本鎖ポリヌクレオチド型があるが、以下、2本鎖ポリヌクレオチドを中心に説明する。

#### 【0048】

配列の設計における好ましい形態は、上記塩基配列設計方法についての説明と同様である。なお、本発明の製造法により製造される2本鎖ポリヌクレオチドはRNAにより構成されることが好ましいが、一部にDNAを含む混成ポリヌクレオチドであってもよい。本明細書では一部にDNAを含むものもsiRNAの概念に含める。また、本発明者らの研究によれば、siRNAは構造・機能的にアシンメトリー性（非対称性）を有する傾向が認められ、RNA干渉を生じさせるという目的からすると、センス鎖の5'末端側の半分、アンチセンス鎖の3'末端側の半分はRNAで構成されることが望ましい。

#### 【0049】

2本鎖ポリヌクレオチドは、一方の鎖が標的遺伝子の塩基配列中に含まれる、(a)から(d)の規則に従う規定配列と相同な塩基配列の3'末端にオーバーハング部を設けて形成され、他方の鎖が、前記規定配列と相同な塩基配列と相補的な配列の3'末端にオーバーハング部を設けて形成される。各鎖の塩基数はオーバーハング部を含め18~24で

あり、より好ましくは20~22、特に好ましくは21である。また、オーバーハング部の塩基数は2であることが好ましい。全体の塩基数が21、そのうちオーバーハング部が塩基数2で構成されるsiRNAは、哺乳類でも細胞毒性を生じさせずに高確率でRNA干渉を生じさせるsiRNAとして好適である。

#### 【0050】

RNAの合成は、例えば、化学合成によって合成してもよいし、また、通常のバイオテクノロジー等の手法に従って行うこともでき、所定の配列を有するDNA鎖を作製し、これを鋳型として転写酵素を用いて一本鎖RNAを合成し、一本鎖RNAを2本鎖化するなどの手法により合成することができる。

#### 【0051】

なお、分子生物学的な基本的手法については、BASIC METHODS IN MOLECULAR BIOLOGY (1986); Sambrookら、MOLECULAR CLONING: A LABORATORY MANUAL、第2版、Cold Spring Harbor Laboratory Press、Cold Spring Harbor、N. Y. (1989)、細胞工学ハンドブック、黒木登志夫ら編、羊土社(1992)、新遺伝子工学ハンドブック改訂第3版、村松ら編、羊土社(1999)など、多くの標準的な実験マニュアルがある。

#### 【0052】

上記本発明の製造方法により得られるポリヌクレオチドとして好ましい形態を示すと、RNA干渉の標的遺伝子の塩基配列中から前記(a)から(d)の規則に従う塩基数13~28の配列部位を検索し、一方の鎖が、標的遺伝子の塩基配列中に含まれる、前記(a)から(d)の規則に従う規定配列と相同な塩基配列の3'末端にオーバーハング部を設けて形成され、他方の鎖が、前記規定配列と相同な塩基配列と相補的な配列の3'末端にオーバーハング部を設けて形成され、各鎖の塩基数が15~30であるように合成された、2本鎖ポリヌクレオチドが例示される。当該ポリヌクレオチドは、RNA干渉を生じさせる確率の高いポリヌクレオチドである。

#### 【0053】

また、siRNAを発現するような発現ベクターを調製することもできる。規定配列を含む配列を発現するベクターを、発現が行われ得る無細胞系または細胞系の条件下に置くことで、発現ベクターを用いて所定のsiRNAを供給することができる。

#### 【0054】

従来、siRNAの設計は、試験者の経験や勘に依存していたため、試行錯誤を繰り返すことが多かった。しかし、本発明の2本鎖ポリヌクレオチドの作製方法により、RNA干渉を生じさせる2本鎖ポリヌクレオチドを高確率で製造することが可能である。上記本発明の検索方法、配列設計方法またはポリヌクレオチドの作製方法によれば、RNA干渉を利用した各種の試験や製造等に要する労力、時間、コストを大幅に削減することが可能である。すなわち、本発明は、遺伝子解析、創薬ターゲットの探索、創薬、遺伝子治療、生物種間の差の研究などのRNA干渉を利用する様々な試験、研究、開発、製造等を大幅に簡便にし、効率の向上を図ることができる。

#### 【0055】

##### <4>遺伝子の発現抑制方法

本発明の遺伝子発現抑制方法は、所定の塩基配列を検索する工程と、検索された塩基配列に基づいてポリヌクレオチドの塩基配列を設計して合成する工程と、得られたポリヌクレオチドを標的遺伝子を含む発現系に導入する工程とを含む。

#### 【0056】

所定の塩基配列を検索する工程は、上記RNA干渉の標的塩基配列検索方法に従う。好ましい態様も上記の通りである。また、検索された塩基配列に基づいてsiRNAの塩基配列を設計して合成する工程は、上記、RNA干渉を生じさせるポリヌクレオチドの塩基配列設計方法、ポリヌクレオチドの作製方法に従って行うことができ、好ましい形態も同様である。



## 【0057】

得られたポリヌクレオチドは、標的遺伝子の発現系に添加して標的遺伝子の発現を抑制する。標的遺伝子の発現系とは、標的遺伝子が発現している体系のことであり、より具体的には、少なくとも標的遺伝子のmRNAが形成される反応系を備える系である。標的遺伝子の発現系としては、In vitro、In vivoのいずれも含まれる。標的遺伝子の発現系として、培養細胞、培養組織、生体などのほか、無細胞系で用いることも可能である。発現抑制をしようとする標的遺伝子（抑制標的遺伝子）は必ずしも検索された配列の由来と一致する生物種のものに限らずともよいが、検索対象遺伝子と抑制対象遺伝子の由来が近縁であればあるほど、特異的かつ効果的に特定遺伝子の抑制を行うことができる。

## 【0058】

標的遺伝子の発現系に導入するとは、標的遺伝子の発現反応系の中に取り込ませるということである。具体的に例を挙げると、標的遺伝子を有する培養細胞に2本鎖ポリヌクレオチドをトランスフェクトし、細胞内に取り込ませる、規定配列およびオーバーハング部からなる塩基配列を有する発現ベクターを作製し、標的遺伝子を有する細胞内に導入する（WO01/36646、WO01/49844）などの手法が挙げられる。

## 【0059】

本発明の遺伝子抑制方法によれば、効率よくRNA干渉を生じさせるポリヌクレオチドを得られるため、効率よく、また簡便に遺伝子を抑制することが可能である。したがって、例えば、対象遺伝子が疾患関連遺伝子の場合、当該疾患関連遺伝子を標的とするsiRNA（またはshRNA）や、このようなsiRNA（またはshRNA）を発現するベクターを、疾患関連遺伝子発現細胞へ導入することにより、疾患関連遺伝子を不活性化することができる。

## 【0060】

## &lt;5&gt; siRNA配列設計プログラム

以下、siRNA配列設計プログラムの実施形態例を説明する。

## 【0061】

## (5-1) プログラムの概要

本プログラムは、ゲノムのシーケンスが進んでいない種、たとえばウマ、ブタなどに対してRNA干渉を行う際に、既に公開されているヒトとマウスのホモログの配列情報をもとに、目的の種で使用可能なsiRNAの配列を計算するものである。本プログラムでsiRNAを設計すれば、標的遺伝子をシーケンスすることなく迅速にRNA干渉を行うことができる。siRNAの設計（計算）にあたっては、GまたはC配分の規則（上記(a)～(d)で示される規則）を考慮してRNA活性を有する確率が高い配列を選ぶとともに、標的遺伝子とは無関係の遺伝子に対してRNA干渉が起こらないようホモロジーサーチによるチェックを行っている。なお、本明細書において「GまたはC」を「G/C」と、「AまたはT」を「A/T」とそれぞれ表記する場合がある。また、「A/T (U)」とは、デオキシリボ核酸による配列の場合T（チミン）であり、リボ核酸による配列の場合U（ウラシル）であることを示す。

## 【0062】

## (5-2) siRNA設計の方針

ヒトの遺伝子Xおよび、そのホモログであるマウスの遺伝子Xの配列が既知であるとする。本プログラムはそれらの配列を読み込み、コード領域（CDS）部分から23塩基以上の完全に共通な配列を探し出す。この共通部分からsiRNAを設計すれば、そのsiRNAはヒトとマウスの遺伝子Xとともに標的とすることができる（第1図）。

## 【0063】

ヒトとマウスとで完全に共通な部分は、他の哺乳類でも保存されている確率が高いと考えられるので、上記のsiRNAはヒトとマウスの遺伝子Xのみならず、他の哺乳類の遺伝子Xに対しても機能することが期待される。つまり、標的遺伝子の配列が未知の動物種であっても、対応するヒトとマウスのホモログについて配列情報が既知であれば、本プロ



グラムを用いて s i RNA を設計することができる。

#### 【0064】

なお、哺乳類においては、実際に機能する s i RNA は配列に規則性があることが判っている（第2図）。本プログラムでは、その規則に適合した配列のみを選ぶようにした。第2図は、RNA i 効果を示す s i RNA の配列の規則性（s i RNA の G / C 配分の規則）を示す図である。第2図において、21塩基の長さをもつ2本のRNA鎖が、それぞれ3'側に2塩基のオーバーハングを持つように5'側の19塩基どうしで塩基対を形成しているような s i RNA で、塩基対を形成している19塩基のうちコーディング側の配列の、1) 3'末端が A / U であること、2) 5'末端が G / C であること、3) 3'側の7文字は、A / U の割合が高いこと、が求められる。特に、1) および 2) の条件は重要である。

#### 【0065】

##### (5-3) プログラムの構造

本プログラムは3つの部分から構成されている。すなわち、(5-3-1) ヒトとマウスとで共通部位の配列（部分配列）を検索する部分、(5-3-2) G / C 配分の規則に基づいて配列にスコアをつける部分、(5-3-3) 無関係の遺伝子を標的にしないようホモロジーサーチによりチェックする部分、である。

#### 【0066】

##### (5-3-1) 共通配列を検索する部分

複数の塩基配列ファイル (file1, file2, file3, ...) を読み込み、全ファイルに共通に出現する23文字の配列をすべて見いだす。

#### 【0067】

##### (計算例)

file1としてヒトの遺伝子FBP1 (NM\_000507:Homo sapiens fructose-1, 6-bisphosphatase 1)、file2としてマウスの遺伝子Fbp1 (NM\_019395:Mus musculus fructose bisphosphatase 1) の配列をプログラムに入力した。その結果、両者の配列（第3図）から両者に共通な23文字の配列（ヒトFBP1とマウスFbp1に共通な配列）が15個見いだされた（第4図）。

#### 【0068】

##### (5-3-2) 配列にスコアをつける部分

前述の G / C 配分の規則に適合した配列のみを選択するために、23文字の配列に対してスコアを付ける。

#### 【0069】

##### (方法)

23文字の配列に対して、以下のようにスコアを付ける。

スコア1: 先頭から21文字目が A / U か。[no=0, yes=1]

スコア2: 先頭から3文字目が G / C か。[no=0, yes=1]

スコア3: 先頭から15文字目から21文字目までの7文字のうち、A / U の数。[0-7]

総合スコア: スコア1~3の積。ただし積が3以下の場合は0とする。

#### 【0070】

##### (計算例)

第4図に示した15個の配列に対して計算を行った結果を、第5図に示す。第5図は、ヒトFBP1とマウスFbp1に共通な配列にスコアを付した図である。なお、第5図に示す配列の後には、スコア1、スコア2、スコア3、総合スコアが順に記載されている。

#### 【0071】

##### (5-3-3) 無関係の遺伝子を標的にしないようチェックする部分

設計した s i RNA が、標的遺伝子とは無関係の遺伝子に対して機能しないよう、公開されているヒト・マウスの全mRNAに対してホモロジーサーチを行い、無関係の遺伝子

がヒットする度合いを評価する。ホモロジーサーチには各種アルゴリズムが使用可能であるが、ここでは、BLASTを使用した例を示す。なお、BLASTを用いる場合は、検索配列が23塩基と短いことを考慮して、Word Sizeを十分に小さくすることが望ましい。

#### 【0072】

BLASTの結果、E valueが10.0以下のヒットのうち、標的遺伝子以外の全てのヒットについて、E valueの逆数の総和を求める（以下、この値をホモロジー・スコアと呼ぶ）。すなわち、ホモロジー・スコア（X）は、下記式により求められる。

#### 【0073】

【数1】

$$X = \sum_{all\ hits} \frac{1}{E}$$

#### 【0074】

注意点：E valueが低いヒットほど、queryの23文字とホモロジーが高く、siRNAの標的とされる危険性が高い。また、ヒットの数が多いほど、より多くの無関係な遺伝子を標的にする確率が高い。この2点を考慮して、siRNAが標的遺伝子と無関係の遺伝子を標的にする危険性を、上の式を用いて評価している。

#### 【0075】

（計算例）

23文字の配列に対して、上記のホモロジーサーチを行った結果と、ホモロジー・スコアを示す（第6図、第7図）。なお、第6図には、ヒトFBP1とマウスFbp1に共通な配列「caccctgaccgcttcgtcatgg」をBLASTした結果が示されており、最初の2行は、マウスFbp1とヒトFBP1がヒットしたものである。ホモロジー・スコアは5.9で、ヒットが少ない例である。この配列のsiRNAは、他の遺伝子を標的にする危険性が低い。また、第7図には、ヒトFBP1とマウスFbp1に共通な配列「gccttctgagaaggatgctctgc」をBLASTした結果が示されている。ヒットが多い例で、ホモロジー・スコアは170.8である。他の遺伝子を標的にする危険性が高いため、siRNAとして適さない。

#### 【0076】

実際には、上記の（5-3-1）、（5-3-2）、（5-3-3）の部分は一体として構成してもよく、第3図に示したヒトとマウスの配列を入力すると、第8図のような出力が直接得られる。ここで、第8図に示す配列の後には、スコア1、スコア2、スコア3、総合スコア、および、ホモロジー・スコアを10倍した値が順に記載されている。なお、処理時間短縮のため、総合スコアが0の配列はホモロジー・スコアを計算しないようにしてもよい。この結果から、siRNAとして「36 caccctgaccgcttcgtcatgg」の部分を使えばよいことがわかる。また、（5-3-1）、（5-3-2）、（5-3-3）の部分のうち1つの機能を単独で利用することもできる。

#### 【0077】

（5-4）実際の計算

ヒト・マウス間のホモログのうち約6400遺伝子のペアに対して、実際に本プログラムでsiRNAの設計をおこなった。その結果、約7割について、ヒトとマウスとに共通な配列で、かつ有効なsiRNAの配列規則性の規則を満たし、無関係な遺伝子を標的としないようなsiRNAが設計できた。

#### 【0078】

これらのsiRNAはヒトとマウスのみならず、多岐の哺乳類で効果的に標的遺伝子を抑える効果があると期待され、家畜や愛玩動物などへの応用など産業的価値が高いと考えられる。また本プログラムを用い、同種内の複数の遺伝子、たとえばEIF2C1とEIF

F2C2を同時に標的とする siRNA を設計することも可能であり、本プログラムが提供する siRNA 設計の手法は、応用範囲が広くきわめて強力なものといえる。さらに、ヒトとマウスの共通な部分の配列を用いて PCR プライマーを設計すれば、多岐の哺乳類で目的の遺伝子を増幅できる、といった利用法もある。

【0079】

なお、siRNA 配列設計プログラムを実行する装置の実施の形態は、下記<7> siRNA 配列設計プログラムを実行する塩基配列処理装置等の欄に詳細に示す。

【0080】

<6> siRNA 配列設計ビジネスモデルシステム

本発明の siRNA 配列設計ビジネスモデルシステムは、siRNA 配列設計プログラムを適用するにあたり、ゲノムデータベース、EST データベース、進化系統樹データベースを本プログラムのロジックに添って単独で或いは組み合わせて参照し、遺伝子配列情報の available 状態に応じた効果的な siRNA を顧客へ提案するというシステムである。available 状態とは、情報が利用可能な状態であることをいう。

【0081】

(1) ゲノム情報が available だが ORF の特定が困難なものについては、EST 情報、等を基にエクソン想定部位に対して効果的な siRNA 候補を抽出し、スプライシングバリエーションを考慮した siRNA 配列とその評価結果を表示する。

【0082】

(2) 遺伝子配列、遺伝子名が明らかなものは、遺伝子配列または遺伝子名を入力後、効果的な siRNA 候補を抽出し、siRNA 配列とその評価結果を表示する。

【0083】

(3) ゲノム情報が available でないものについては、同種の遺伝子機能を保存している類縁関係（同属、乃至は起源を同じとする）にある生物種の遺伝子配列、或いは進化系統樹的に挟まれた、ゲノム配列が available な 2 種類以上の生物種の遺伝子配列、を用いて効果的な siRNA 候補を抽出し、siRNA 配列とその評価結果を表示する。

【0084】

(4) 感染症の遺伝子機能解析、創薬ターゲット発掘には、微生物のゲノムデータベース・進化系統樹データベースへ更に微生物のアポトーシス誘導部位情報、機能発現部位情報を組み合わせて網羅的な siRNA 候補配列を求める手法が有効である。

【0085】

<7> siRNA 配列設計プログラムを実行する塩基配列処理装置等

以下、上述した siRNA 配列設計プログラムを実行する装置である、本発明にかかる塩基配列処理装置、塩基配列処理方法をコンピュータに実行させるプログラム、記録媒体、および、塩基配列処理システムの実施の形態を図面に基づいて詳細に説明する。なお、この実施の形態によりこの発明が限定されるものではない。

【0086】

[本発明の概要]

以下、本発明の概要について説明し、その後、本発明の構成および処理等について詳細に説明する。第 12 図は、本発明の基本原理を示す原理構成図である。

【0087】

本発明は、概略的に、以下の基本的特徴を有する。すなわち、本発明は、RNA 干渉の標的遺伝子の塩基配列情報を取得して、塩基配列情報の予め定めた塩基数の配列部位である部分塩基配列情報を作成する（ステップ S-1）。

【0088】

ここで、ステップ S-1 において、塩基配列情報の標的遺伝子のコード領域、または、転写領域に対応する部位から、予め定めた塩基数の部分塩基配列情報を作成してもよい。また、異なる生物に由来する複数の塩基配列情報（例えば、ヒトの塩基配列情報およびマウスの塩基配列情報など）の間で共通する、予め定めた塩基数の部分塩基配列情報を作成

してもよい。また、同じ生物種における類似する複数の塩基配列情報の間で共通する、予め定めた塩基数の共通な部分塩基配列情報を作成してもよい。また、異なる生物に由来する複数の塩基配列情報の標的遺伝子のコード領域、または、転写領域に対応する部位から、予め定めた塩基数の共通な部分塩基配列情報を作成してもよい。また、同じ生物種における類似する複数の塩基配列情報の標的遺伝子のコード領域、または、転写領域に対応する部位から、予め定めた塩基数の共通な部分塩基配列情報を作成してもよい。これにより、標的遺伝子に特異的にRNA干渉を生じさせる規定配列を効率よく選択することができ、計算負荷を軽減できる。

#### 【0089】

さらに、ステップS-1において、オーバーハング部位を含む部分塩基配列情報を作成してもよい。具体的には、例えば、オーバーハング部位が含まれていることを表すオーバーハング部位含有情報が付加された部分塩基配列情報を作成してもよい。すなわち、部分塩基配列情報とオーバーハング部位含有情報とを相互に関連付けてもよい。これにより、最初から規定配列およびオーバーハング部位を含めて選択を行い、設計することができるようになる。

#### 【0090】

なお、上述の予め定めた塩基数の上限は、オーバーハング部位を含まない場合は、好ましくは28以下、より好ましくは22以下、さらに好ましくは20以下であり、オーバーハング部位を含む場合は、好ましくは32以下、より好ましくは26以下、さらに好ましくは24以下である。また、予め定めた塩基数の下限は、オーバーハング部位を含まない場合は、好ましくは13以上、より好ましくは16以上、さらに好ましくは18以上であり、オーバーハング部位を含む場合は、好ましくは17以上、より好ましくは20以上、さらに好ましくは22以上である。そして、最も好ましい予め定めた塩基数は、オーバーハング部位を含まない場合は、19であり、オーバーハング部位を含む場合は、23である。これにより、哺乳類においても細胞毒性を生じさせずにRNA干渉を生じさせる規定配列を効率よく選択することができる。

#### 【0091】

ついで、ステップS-1にて作成された部分塩基配列情報の3'末端の塩基がアデニン、チミン、または、ウラシルであるか否かを判定する(ステップS-2)。なお、具体的には、例えば、3'末端の塩基がアデニン、チミン、または、ウラシルであるときには「1」を、そうでないときには「0」を判定結果として出力してもよい。

#### 【0092】

ついで、ステップS-1にて作成された部分塩基配列情報の5'末端の塩基がグアニン、または、シトシンであるか否かを判定する(ステップS-3)。なお、具体的には、例えば、5'末端の塩基がグアニン、または、シトシンであるときには「1」を、そうでないときには「0」を判定結果として出力してもよい。

#### 【0093】

ついで、ステップS-1にて作成された部分塩基配列情報の3'末端の7塩基からなる塩基配列情報が、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であるか否かを判定する(ステップS-4)。なお、具体的には、例えば、部分塩基配列情報の3'末端の7塩基からなる塩基配列情報に含まれる、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基の数を判定結果として出力してもよい。ステップS-4における判定の規則は、ステップS-1にて作成された部分塩基配列情報の3'末端近傍の塩基配列情報がアデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であることを規定しており、具体的に検索を行う際の一指標として3'末端から7塩基の範囲内の塩基配列情報がアデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であることを規定している。

#### 【0094】

ここで、ステップS-4において、「リッチな塩基配列情報」は、上述の<1>RNA干渉の標的塩基配列検索方法に記述されている「リッチな配列」であり、具体的には、例えば、ステップS-1にて作成された部分塩基配列情報が19塩基程度からなる場合、当該部分塩基配列情報の7塩基からなる塩基配列情報が、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基が少なくとも好ましくは3塩基以上、より好ましくは4塩基以上、特に好ましくは5塩基以上含まれる塩基配列情報である。

#### 【0095】

また、ステップS-2からステップS-4において、オーバーハング部位を含む部分塩基配列情報を判定する場合、部分塩基配列情報のオーバーハング部位を除いた配列部位を判定対象として判定する。

#### 【0096】

ついで、ステップS-2、ステップS-3、および、ステップS-4にて判定された結果に基づいて、ステップS-1にて作成された部分塩基配列情報から標的遺伝子に特異的にRNA干渉を生じさせる規定配列情報を選択する（ステップS-5）。

#### 【0097】

具体的には、例えば、ステップS-2にて3'末端の塩基がアデニン、チミン、または、ウラシルであると判定され、ステップS-3にて5'末端の塩基がグアニン、または、シトシンであると判定され、かつ、ステップS-4にて部分塩基配列情報の3'末端の7塩基からなる塩基配列情報が、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であると判定された部分塩基配列情報を規定配列情報として選択する。ここで、具体的には、例えば、ステップS-2、ステップS-3、および、ステップS-4にて出力される数値の積を算出し、当該積の値に基づいて、ステップS-1にて作成された部分塩基配列情報から規定配列情報を選択してもよい。

#### 【0098】

これにより、哺乳類などにおいて、RNA干渉を生じさせる確率が極めて高い、つまりRNA干渉に有効なsiRNAの配列を効率よく、容易に作成することができる。

#### 【0099】

ここで、ステップS-5にて選択された規定配列情報の少なくとも一方の末端にオーバーハング部位を付加してもよい。なお、例えば、ターゲット側を検索する場合は、規定配列情報の両末端にオーバーハング部位を付加してもよい。これにより、RNA干渉を生じさせるポリヌクレオチドの設計を簡便化することができる。

#### 【0100】

なお、上述のオーバーハング部位の塩基数は、上述の<2>RNA干渉を生じさせるポリヌクレオチドの塩基配列設計方法に記述されている塩基数であり、具体的には、例えば、2が特に好適である。

#### 【0101】

また、ステップS-5にて選択された規定配列情報と同一または類似の塩基配列情報を、他の塩基配列情報（例えばNCBIのRefSeq（Reference Sequence project）などの公的データベースで公開されている塩基配列情報）から、例えばBLAST、FASTA、sssearchなどの既存のホモロジー検索手法を用いて検索し、検索された同一または類似の塩基配列情報に基づいて、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にするか否かを評価してもよい。

#### 【0102】

具体的には、例えば、ステップS-5にて選択された規定配列情報と同一または類似の塩基配列情報を、他の塩基配列情報（例えばNCBIのRefSeqなどの公的データベースで公開されている塩基配列情報）から、例えばBLAST、FASTA、sssearchなどの既存のホモロジー検索手法を用いて検索し、検索された同一または類似の塩基配列情報のうち標的遺伝子とは無関係の遺伝子の塩基配列情報の総数、および、標的遺伝

子とは無関係の遺伝子の塩基配列情報に付された同一類似の度合いを示す値（例えばBLAST、FASTAおよびsssearchの場合は、「E value」）に基づいて、同一類似の度合いを示す値の逆数の総和を算出し、算出された総和に基づいて（例えば、算出された総和の大小などに基づいて）、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にするか否かを評価してもよい。

#### 【0103】

これにより、規定配列情報から、標的とする遺伝子のみの特異的にRNA干渉を生じさせる配列を選出することができる。

#### 【0104】

以上、本発明により選択された、標的遺伝子と関係のない遺伝子についてまでRNA干渉を生じさせない規定配列情報に基づいてRNAの合成を行えば、それにかかる労力、時間、費用を従来に比べ大幅に軽減することができる。

#### 【0105】

##### [システム構成]

まず、本システムの構成について説明する。第13図は、本発明が適用される本システムの構成の一例を示すブロック図であり、該構成のうち本発明に関係する部分のみを概念的に示している。

#### 【0106】

本システムは、概略的に、RNA干渉の標的遺伝子の塩基配列情報を処理する塩基配列処理装置100と、配列情報や構造情報等に関する外部データベースやホモロジー検索等の外部プログラム等を提供する外部システム200とを、ネットワーク300を介して通信可能に接続して構成されている。

#### 【0107】

第13図において、ネットワーク300は、塩基配列処理装置100と外部システム200とを相互に接続する機能を有し、例えば、インターネット等である。

#### 【0108】

第13図において、外部システム200は、ネットワーク300を介して、塩基配列処理装置100と相互に接続され、利用者に対して配列情報や構造情報等に関する外部データベースやホモロジー検索やモチーフ検索等の外部プログラムを実行するウェブサイトを提供する機能を有する。

#### 【0109】

ここで、外部システム200は、WEBサーバやASPサーバ等として構成してもよく、そのハードウェア構成は、一般に市販されるワークステーション、パーソナルコンピュータ等の情報処理装置およびその付属装置により構成してもよい。また、外部システム200の各機能は、外部システム200のハードウェア構成中のCPU、ディスク装置、メモリ装置、入力装置、出力装置、通信制御装置等およびそれらを制御するプログラム等により実現される。

#### 【0110】

第13図において塩基配列処理装置100は、概略的に、塩基配列処理装置100の全体を統括的に制御するCPU等の制御部102、通信回線等に接続されるルータ等の通信装置（図示せず）に接続される通信制御インターフェース部104、入力装置112や出力装置114に接続される入出力制御インターフェース部108、および、各種のデータベースやテーブルなどを格納する記憶部106を備えて構成されており、これら各部は任意の通信路を介して通信可能に接続されている。さらに、この塩基配列処理装置100は、ルータ等の通信装置および専用線等の有線または無線の通信回線を介して、ネットワーク300に通信可能に接続されている。

#### 【0111】

記憶部106に格納される各種のデータベースやテーブル（標的遺伝子塩基配列ファイル106a～標的遺伝子アノテーションデータベース106h）は、固定ディスク装置等のストレージ手段であり、各種処理に用いる各種のプログラムやテーブルやファイルやデ

データベースやウェブページ用ファイル等を格納する。

【0112】

これら記憶部106の各構成要素のうち、標的遺伝子塩基配列ファイル106aは、RNA干渉の標的遺伝子の塩基配列情報を格納する標的遺伝子塩基配列格納手段である。第14図は、標的遺伝子塩基配列ファイル106aに格納される情報の一例を示す図である。

【0113】

標的遺伝子塩基配列ファイル106aに格納される情報は、第14図に示すように、RNA干渉の標的遺伝子の塩基配列情報を一意に識別する塩基配列識別情報（例えば、第14図の「NM\_000507」）と、塩基配列情報（例えば、第14図の「ATGGCTGA...AGTGA」）とを相互に関連付けて構成されている。

【0114】

また、部分塩基配列ファイル106bは、RNA干渉の標的遺伝子の塩基配列情報の予め定めた塩基数の配列部位である部分塩基配列情報を格納する部分塩基配列格納手段である。第15図は、部分塩基配列ファイル106bに格納される情報の一例を示す図である。

【0115】

部分塩基配列ファイル106bに格納される情報は、第15図に示すように、部分塩基配列情報を一意に識別する部分塩基配列識別情報（例えば、第15図の「NM\_000507:36」）と、部分塩基配列情報（例えば、第15図の「caccct...tcattgg」）と、オーバーハング部位が含まれていることを表すオーバーハング部位含有情報（例えば、第15図の「含有」）とを相互に関連付けて構成されている。

【0116】

また、判定結果ファイル106cは、後述する3'末端塩基判定部102b、5'末端塩基判定部102c、および、特定塩基含有判定部102dにて判定された結果を格納する判定結果手段である。第16図は、判定結果ファイル106cに格納される情報の一例を示す図である。

【0117】

判定結果ファイル106cに格納される情報は、第16図に示すように、部分塩基配列識別情報（例えば、第16図の「NM\_000507:36」）と、3'末端塩基判定部102bにて判定された結果である3'末端塩基判定結果（例えば、第16図の「1」）と、5'末端塩基判定部102cにて判定された結果である5'末端塩基判定結果（例えば、第16図の「1」）と、特定塩基含有判定部102dにて判定された結果である特定塩基含有判定結果（例えば、第16図の「4」）と、3'末端塩基判定部102b、5'末端塩基判定部102c、および、特定塩基含有判定部102dにて判定された結果を総合した結果である総合判定結果（例えば、第16図の「4」）とを相互に関連付けて構成されている。

【0118】

なお、第16図において、3'末端塩基判定結果および5'末端塩基判定結果は、3'末端塩基判定部102b、5'末端塩基判定部102cのそれぞれにて「含む」と判定された場合「1」を設定し、「含まない」と判定された場合「0」を設定した場合の一例である。また、第16図において、特定塩基含有判定結果は、部分塩基配列情報の3'末端の7塩基からなる塩基配列情報に含まれる、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基数を設定した場合の一例である。さらに、第16図において、総合判定結果は、3'末端塩基判定結果、5'末端塩基判定結果、および、特定塩基含有判定結果の積を設定した場合の一例である。なお、具体的には、例えば、当該積が3以下の場合は、「0」を設定してもよい。

【0119】

また、規定配列ファイル106dは、標的遺伝子に特異的にRNA干渉を生じさせる部分塩基配列情報である規定配列情報を格納する規定配列格納手段である。第17図は、規



定配列ファイル106dに格納される情報の一例を示す図である。

【0120】

規定配列ファイル106dに格納される情報は、第17図に示すように、部分塩基配列識別情報（例えば、第17図の「NM\_000507:36」）と、標的遺伝子に特異的にRNA干渉を生じさせる部分塩基配列情報である規定配列情報（例えば、第17図の「caccct...tcatgg」）とを相互に関連付けて構成されている。

【0121】

また、参照配列データベース106eは、後述する同一類似塩基配列検索部102gにて規定配列情報と同一または類似の塩基配列情報を検索するために参照する塩基配列情報である参照塩基配列情報を格納するデータベースである。なお、参照配列データベース106eは、インターネットを経由してアクセスする外部の塩基配列情報データベースであってもよく、また、これらのデータベースをコピーしたり、オリジナルの配列情報を格納したり、さらに独自のアノテーション情報等を付加したりして作成したインハウスデータベースであってもよい。第18図は、参照配列データベース106eに格納される情報の一例を示す図である。

【0122】

参照配列データベース106eに格納される情報は、第18図に示すように、参照配列識別情報（例えば、第18図の「ref|NM\_015820.1|」）と、参照塩基配列情報（例えば、第18図の「caccct...gcatgg」）とを相互に関連付けて構成されている。

【0123】

また、同一類似度ファイル106fは、後述する同一類似塩基配列検索部102gにて検索された同一または類似の塩基配列情報に付された同一類似の度合いを示す値である同一類似度を格納する同一類似度格納手段である。第19図は、同一類似度ファイル106fに格納される情報の一例を示す図である。

【0124】

同一類似度ファイル106fに格納される情報は、第19図に示すように、部分塩基配列識別情報（例えば、第19図の「NM\_000507:36」）と、参照配列識別情報（例えば、第19図の「ref|NM\_015820.1|」および「ref|NM\_003837.1|」）と、同一類似度（例えば、第19図の「0.52」）とを相互に関連付けて構成されている。

【0125】

また、評価結果ファイル106gは、後述する無関係遺伝子標的評価部102hにて標的遺伝子とは無関係の遺伝子を標的にするか否かを評価した結果を格納する評価結果格納手段である。第20図は、評価結果ファイル106gに格納される情報の一例を示す図である。

【0126】

評価結果ファイル106gに格納される情報は、第20図に示すように、部分塩基配列識別情報（例えば、第20図の「NM\_000507:36」および「NM\_000507:441」）と、後述する総和算出部102mにて算出された総和（例えば、第20図の「5.9」および「170.8」）と、評価結果（例えば、第20図の「非標的」および「標的」）とを相互に関連付けて構成されている。なお、第20図において、「非標的」は、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にしないことを意味し、また、「標的」は、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にすることを意味する。

【0127】

また、標的遺伝子アノテーションデータベース106hは、標的遺伝子に関するアノテーション情報を格納する標的遺伝子アノテーション格納手段である。なお、標的遺伝子アノテーションデータベース106hは、インターネットを経由してアクセスする、遺伝子に関するアノテーション情報を格納する外部のアノテーションデータベースであってもよ



く、また、これらのデータベースをコピーしたり、オリジナルの配列情報を格納したり、さらに独自のアノテーション情報等を付加したりして作成したインハウスデータベースであってもよい。

#### 【0128】

標的遺伝子アノテーションデータベース106hに格納される情報は、標的遺伝子を識別する標的遺伝子識別情報（例えば標的遺伝子の遺伝子名、アクセッション（Accession）番号など（例えば、第3図の最上部に記載の「NM\_000507」、「FBP1」））と標的遺伝子に関する簡略的な情報（例えば第3図の最上部に記載の「Homo sapiens fructose-1, 6-bisphosphatase 1」）とを相互に関連付けて構成されている。

#### 【0129】

また、第13図において、通信制御インターフェース部104は、塩基配列処理装置100とネットワーク300（またはルータ等の通信装置）との間における通信制御を行う。すなわち、通信制御インターフェース部104は、他の端末と通信回線を介してデータを通信する機能を有する。

#### 【0130】

また、第13図において、入出力制御インターフェース部108は、入力装置112や出力装置114の制御を行う。ここで、出力装置114としては、モニタ（家庭用テレビを含む）の他、スピーカを用いることができる（なお、以下においては出力装置114をモニタとして記載する場合がある）。また、入力装置112としては、キーボード、マウス、および、マイク等を用いることができる。また、モニタも、マウスと協働してポインティングデバイス機能を実現する。

#### 【0131】

また、第13図において、制御部102は、OS（Operating System）等の制御プログラム、各種の処理手順等を規定したプログラム、および所要データを格納するための内部メモリを有し、これらのプログラム等により、種々の処理を実行するための情報処理を行う。制御部102は、機能概念的に、部分塩基配列作成部102a、3'末端塩基判定部102b、5'末端塩基判定部102c、特定塩基含有判定部102d、規定配列選択部102e、オーバーハング部位付加部102f、同一類似塩基配列検索部102g、および、無関係遺伝子標的評価部102hを備えて構成されている。

#### 【0132】

このうち、部分塩基配列作成部102aは、RNA干渉の標的遺伝子の塩基配列情報を取得して、塩基配列情報の予め定めた塩基数の配列部位である部分塩基配列情報を作成する部分塩基配列作成手段である。ここで、部分塩基配列作成部102aは、第21図に示すように、領域指定塩基配列作成部102i、共通塩基配列作成部102j、および、オーバーハング部位含有塩基配列作成部102kをさらに含んで構成される。

#### 【0133】

第21図は、本発明が適用される本システムの部分塩基配列作成部102aの構成の一例を示すブロック図であり、該構成のうち本発明に係る部分のみを概念的に示している。

#### 【0134】

第21図において、領域指定塩基配列作成部102iは、塩基配列情報の標的遺伝子のコード領域、または、転写領域に対応する部位から、予め定めた塩基数の部分塩基配列情報を作成する領域指定塩基配列作成手段である。

#### 【0135】

また、共通塩基配列作成部102jは、異なる生物に由来する複数の塩基配列情報の間で共通する、予め定めた塩基数の部分塩基配列情報を作成する共通塩基配列作成手段である。

#### 【0136】

また、オーバーハング部位含有塩基配列作成部102kは、オーバーハング部位を含む

部分塩基配列情報を作成するオーバーハング部位含有塩基配列作成手段である。

【0137】

再び第13図に戻り、3'末端塩基判定部102bは、部分塩基配列情報の3'末端の塩基がアデニン、チミン、または、ウラシルであるか否かを判定する3'末端塩基判定手段である。

【0138】

また、5'末端塩基判定部102cは、部分塩基配列情報の5'末端の塩基がグアニン、または、シトシンであるか否かを判定する5'末端塩基判定手段である。

【0139】

また、特定塩基含有判定部102dは、部分塩基配列情報の3'末端の7塩基からなる塩基配列情報が、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であるか否かを判定する特定塩基含有判定手段である。

【0140】

また、規定配列選択部102eは、3'末端塩基判定部102b、5'末端塩基判定部102c、および、特定塩基含有判定部102cにて判定された結果に基づいて、部分塩基配列情報から標的遺伝子に特異的にRNA干渉を生じさせる規定配列情報を選択する規定配列選択手段である。

【0141】

また、オーバーハング部位付加部102fは、規定配列情報の少なくとも一方の末端にオーバーハング部位を付加するオーバーハング部位付加手段である。

【0142】

また、同一類似塩基配列検索部102gは、規定配列情報と同一または類似の塩基配列情報を他の塩基配列情報から検索する同一類似塩基配列検索手段である。

【0143】

また、無関係遺伝子標的評価部102hは、同一または類似の塩基配列情報に基づいて、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にするか否かを評価する無関係遺伝子標的評価手段である。ここで、無関係遺伝子標的評価部102hは、第22図に示すように、総和算出部102m、および、総和基準評価部102nをさらに含んで構成される。

【0144】

第22図は、本発明が適用される本システムの無関係遺伝子標的評価部102hの構成の一例を示すブロック図であり、該構成のうち本発明に関する部分のみを概念的に示している。

【0145】

第22図において、総和算出部102mは、同一または類似の塩基配列情報のうち標的遺伝子とは無関係の遺伝子の塩基配列情報の総数、および、標的遺伝子とは無関係の遺伝子の塩基配列情報に付された同一類似の度合いを示す値（同一類似度）に基づいて、同一類似の度合いを示す値の逆数の総和を算出する総和算出手段である。

【0146】

また、総和基準評価部102nは、総和算出部102mにて算出された総和に基づいて、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にするか否かを評価する総和基準標的評価手段である。

【0147】

なお、これら各部によって行なわれる処理の詳細については、後述する。

【0148】

[システムの処理]

次に、このように構成された本実施の形態における本システムの処理の一例について、以下に第23図、および、第24図を参照して詳細に説明する。

【0149】

## 【メイン処理】

まず、メイン処理の詳細について第23図等を参照して説明する。第23図は、本実施形態における本システムのメイン処理の一例を示すフローチャートである。

## 【0150】

まず、塩基配列処理装置100は、部分塩基配列作成部102aにて行われる部分塩基配列作成処理により、RNA干渉の標的遺伝子の塩基配列情報を取得して標的遺伝子塩基配列ファイル106aの所定の記憶領域に格納し、塩基配列情報の予め定めた塩基数の配列部位である部分塩基配列情報を作成し、作成された部分塩基配列情報を部分塩基配列ファイル106bの所定の記憶領域に格納する（ステップSA-1）。

## 【0151】

ここで、ステップSA-1において、部分塩基配列作成部102aは、領域指定塩基配列作成部102iの処理により、塩基配列情報の標的遺伝子のコード領域、または、転写領域に対応する部位から、予め定めた塩基数の部分塩基配列情報を作成し、作成された部分塩基配列情報を部分塩基配列ファイル106bの所定の記憶領域に格納してもよい。

## 【0152】

また、ステップSA-1において、部分塩基配列作成部102aは、共通塩基配列作成部102jの処理により、異なる生物に由来する複数の塩基配列情報（例えば、ヒトの塩基配列情報およびマウスの塩基配列情報など）の間で共通する、予め定めた塩基数の部分塩基配列情報を作成し、作成された部分塩基配列情報を部分塩基配列ファイル106bの所定の記憶領域に格納してもよい。なお、同じ生物種における類似する複数の塩基配列情報の間で共通する、予め定めた塩基数の共通な部分塩基配列情報を作成してもよい。

## 【0153】

また、ステップSA-1において、部分塩基配列作成部102aは、領域指定塩基配列作成部102i、および、共通塩基配列作成部102jの処理により、異なる生物に由来する複数の塩基配列情報の標的遺伝子のコード領域、または、転写領域に対応する部位から、予め定めた塩基数の共通な部分塩基配列情報を作成し、作成された部分塩基配列情報を部分塩基配列ファイル106bの所定の記憶領域に格納してもよい。なお、同じ生物種における類似する複数の塩基配列情報の標的遺伝子のコード領域、または、転写領域に対応する部位から、予め定めた塩基数の共通な部分塩基配列情報を作成してもよい。

## 【0154】

さらに、ステップSA-1において、部分塩基配列作成部102aは、オーバーハング部位含有塩基配列作成部102kの処理により、オーバーハング部位を含む部分塩基配列情報を作成してもよい。具体的には、例えば、部分塩基配列作成部102aは、オーバーハング部位含有塩基配列作成部102kの処理により、オーバーハング部位が含まれていることを表すオーバーハング部位含有情報が付加された部分塩基配列情報を作成し、作成された部分塩基配列情報とオーバーハング部位含有情報と相互に関連付けて部分塩基配列ファイル106bの所定の記憶領域に格納してもよい。

## 【0155】

なお、上述の予め定めた塩基数の上限は、オーバーハング部位を含まない場合は、好ましくは28以下、より好ましくは22以下、さらに好ましくは20以下であり、オーバーハング部位を含む場合は、好ましくは32以下、より好ましくは26以下、さらに好ましくは24以下である。また、予め定めた塩基数の下限は、オーバーハング部位を含まない場合は、好ましくは13以上、より好ましくは16以上、さらに好ましくは18以上であり、オーバーハング部位を含む場合は、好ましくは17以上、より好ましくは20以上、さらに好ましくは22以上である。そして、最も好ましい予め定めた塩基数は、オーバーハング部位を含まない場合は、19であり、オーバーハング部位を含む場合は、23である。

## 【0156】

ついで、塩基配列処理装置100は、3'末端塩基判定部102bの処理により、ステップSA-1にて作成された部分塩基配列情報の3'末端の塩基がアデニン、チミン、ま

たは、ウラシルであるか否かを判定し、判定した結果を判定結果ファイル106cの所定の記憶領域に格納する（ステップSA-2）。具体的には、例えば、塩基配列処理装置100は、3'末端塩基判定部102bの処理により、ステップSA-1にて作成された部分塩基配列情報の3'末端の塩基がアデニン、チミン、または、ウラシルであるときには「1」を、そうでないときには「0」を判定結果ファイル106cの所定の記憶領域に格納してもよい。

#### 【0157】

ついで、塩基配列処理装置100は、5'末端塩基判定部102cの処理により、ステップSA-1にて作成された部分塩基配列情報の5'末端の塩基がグアニン、または、シトシンであるか否かを判定し、判定した結果を判定結果ファイル106cの所定の記憶領域に格納する（ステップSA-3）。具体的には、例えば、塩基配列処理装置100は、5'末端塩基判定部102cの処理により、ステップSA-1にて作成された部分塩基配列情報の5'末端の塩基がグアニン、または、シトシンであるときには「1」を、そうでないときには「0」を判定結果ファイル106cの所定の記憶領域に格納してもよい。

#### 【0158】

ついで、塩基配列処理装置100は、特定塩基含有判定部102dの処理により、ステップSA-1にて作成された部分塩基配列情報の3'末端の7塩基からなる塩基配列情報が、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であるか否かを判定し、判定した結果を判定結果ファイル106cの所定の記憶領域に格納する（ステップSA-4）。具体的には、例えば、塩基配列処理装置100は、特定塩基含有判定部102dの処理により、ステップSA-1にて作成された部分塩基配列情報の3'末端の7塩基からなる塩基配列情報に含まれる、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基の数を判定結果ファイル106cの所定の記憶領域に格納してもよい。ステップSA-4における判定の規則は、ステップSA-1にて作成された部分塩基配列情報の3'末端近傍の塩基配列情報がアデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であることを規定しており、具体的に検索を行う際の一指標として3'末端から7塩基の範囲内の塩基配列情報がアデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であることを規定している。

#### 【0159】

ここで、ステップSA-4において、「リッチな塩基配列情報」は、上述の<1>RNA干渉の標的塩基配列検索方法に記述されている「リッチな配列」であり、具体的には、例えば、ステップSA-1にて作成された部分塩基配列情報が19塩基程度からなる場合、当該部分塩基配列情報の3'末端の7塩基からなる塩基配列情報が、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基が少なくとも好ましくは3塩基以上、より好ましくは4塩基以上、特に好ましくは5塩基以上含まれる塩基配列情報である。

#### 【0160】

また、ステップSA-2からステップSA-4において、オーバーハング部位を含む部分塩基配列情報を判定する場合、部分塩基配列情報のオーバーハング部位を除いた配列部位を判定対象として判定する。

#### 【0161】

ついで、塩基配列処理装置100は、規定配列選択部102eの処理により、ステップSA-2、ステップSA-3、および、ステップSA-4にて判定された結果に基づいて、ステップSA-1にて作成された部分塩基配列情報から標的遺伝子に特異的にRNA干渉を生じさせる規定配列情報を選択し、規定配列ファイル106dの所定の記憶領域に格納する（ステップSA-5）。

#### 【0162】

具体的には、例えば、塩基配列処理装置100は、規定配列選択部102eの処理によ

り、ステップSA-2にて3'末端の塩基がアデニン、チミン、または、ウラシルであると判定され、ステップSA-3にて5'末端の塩基がグアニン、または、シトシンであると判定され、かつ、ステップSA-4にて部分塩基配列情報の3'末端の7塩基からなる塩基配列情報が、アデニン、チミン、および、ウラシルからなる群より選ばれる1種または2種以上の塩基がリッチな塩基配列情報であると判定された部分塩基配列情報を規定配列情報として選択し、規定配列ファイル106dの所定の記憶領域に格納する。ここで、具体的には、例えば、塩基配列処理装置100は、規定配列選択部102eの処理により、ステップSA-2、ステップSA-3、および、ステップSA-4にて出力される数値の積を算出し、当該積の値に基づいて、ステップSA-1にて作成された部分塩基配列情報から規定配列情報を選択してもよい。

#### 【0163】

ここで、塩基配列処理装置100は、オーバーハング部位付加部102fの処理により、ステップSA-5にて選択された規定配列情報の少なくとも一方の末端にオーバーハング部位を付加して、規定配列ファイル106dの所定の記憶領域に格納してもよい。具体的には、例えば、塩基配列処理装置100は、オーバーハング部位付加部102fの処理により、規定配列ファイル106dの規定配列情報の項に記憶されている規定配列情報を、少なくとも一方の末端にオーバーハング部位が付加された規定配列情報に書き換えてもよい。なお、例えば、ターゲット側を検索する場合は、規定配列情報の両末端にオーバーハング部位を付加してもよい。

#### 【0164】

なお、上述のオーバーハング部位の塩基数は、上述の<2>RNA干渉を生じさせるポリヌクレオチドの塩基配列設計方法に記述されている塩基数であり、具体的には、例えば、2が特に好適である。

#### 【0165】

また、塩基配列処理装置100は、同一類似塩基配列検索部102gの処理により、ステップSA-5にて選択された規定配列情報と同一または類似の塩基配列情報を、他の塩基配列情報（例えばNCBIのRefSeqなどの公的データベースで公開されている塩基配列情報）から、例えばBLAST、FASTA、searchなどの既存のホモロジー検索手法を用いて検索し、無関係遺伝子標的評価部102hにて行われる無関係遺伝子標的評価処理により、検索された同一または類似の塩基配列情報に基づいて、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にするか否かを評価してもよい。

#### 【0166】

具体的には、例えば、塩基配列処理装置100は、同一類似塩基配列検索部102gの処理により、ステップSA-5にて選択された規定配列情報と同一または類似の塩基配列情報を、他の塩基配列情報（例えばNCBIのRefSeqなどの公的データベースで公開されている塩基配列情報）から、例えばBLAST、FASTA、searchなどの既存のホモロジー検索手法を用いて検索し、無関係遺伝子標的評価部102hは、総和算出部102mの処理により、検索された同一または類似の塩基配列情報のうち標的遺伝子とは無関係の遺伝子の塩基配列情報の総数、および、標的遺伝子とは無関係の遺伝子の塩基配列情報に付された同一類似の度合いを示す値（例えばBLAST、FASTAおよびsearchの場合は、「E value」）に基づいて、同一類似の度合いを示す値の逆数の総和を算出し、無関係遺伝子標的評価部102hは、総和基準評価部102nの処理により、算出された総和に基づいて、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にするか否かを評価してもよい。

#### 【0167】

ここで、無関係遺伝子標的評価部102hにて行われる無関係遺伝子標的評価処理の詳細について第24図を参照して説明する。

#### 【0168】

第24図は、本実施形態における本システムの無関係遺伝子標的評価処理の一例を示すフローチャートである。

## 【0169】

まず、塩基配列処理装置100は、同一類似塩基配列検索部102gの処理により、ステップSA-5にて選択された規定配列情報と同一または類似の塩基配列情報を、他の塩基配列情報（例えばNCBIのRefSeqなどの公的データベースで公開されている塩基配列情報）から、例えばBLAST、FASTA、sssearchなどの既存のホモロジー検索手法を用いて検索し、規定配列情報の識別情報（第19図における「部分塩基配列識別情報」）と、検索された同一または類似の塩基配列情報の識別情報（第19図における「参照配列識別情報」）と、検索された同一または類似の塩基配列情報に付された同一類似の度合いを示す値（例えばBLAST、FASTAおよびsssearchの場合は、「E value」）（第19図における「同一類似度」）とを相互に関連付けて、同一類似度ファイル106fの所定の記憶領域に格納する。

## 【0170】

ついで、無関係遺伝子標的評価部102hは、総和算出部102mの処理により、検索された同一または類似の塩基配列情報のうち標的遺伝子とは無関係の遺伝子の塩基配列情報の総数、および、標的遺伝子とは無関係の遺伝子の塩基配列情報に付された同一類似の度合いを示す値（例えばBLAST、FASTAおよびsssearchの場合は、「E value」）に基づいて、同一類似の度合いを示す値の逆数の総和を算出し、規定配列情報の識別情報（第20図における「部分塩基配列識別情報」）と算出された総和（第20図における「総和」）とを相互に関連付けて、評価結果ファイル106gの所定の記憶領域に格納する（ステップSB-1）。

## 【0171】

ついで、無関係遺伝子標的評価部102hは、総和基準評価部102nの処理により、ステップSB-1にて算出された総和に基づいて（例えば、ステップSB-1にて算出された総和の大小などに基づいて）、規定配列情報が標的遺伝子とは無関係の遺伝子を標的にするか否かを評価し、評価結果（第20図における「非標的」および「標的」）を評価結果ファイル106gの所定の記憶領域に格納する（ステップSB-2）。

## 【0172】

これにて、メイン処理が終了する。

## 【0173】

〔他の実施の形態〕

さて、これまで本発明の実施の形態について説明したが、本発明は、上述した実施の形態以外にも、請求の範囲に記載した技術的思想の範囲内において種々の異なる実施の形態にて実施されてよいものである。

## 【0174】

例えば、塩基配列処理装置100がスタンドアローンの形態で処理を行う場合を一例に説明したが、塩基配列処理装置100とは別筐体で構成されるクライアント端末からの要求に応じて処理を行い、その処理結果を当該クライアント端末に返却するように構成してもよい。具体的には、例えば、クライアント端末はRNA干渉の標的遺伝子の名称（例えば遺伝子名、アクセッション番号など）または標的遺伝子に係る塩基配列情報を塩基配列処理装置100に送信し、塩基配列処理装置100は当該名称に対応する塩基配列情報またはクライアント端末より送信された塩基配列情報に対して制御部102で行われる上述した各処理を行い、標的遺伝子に特異的にRNA干渉を生じさせる規定配列情報を選択してクライアント端末に送信してもよい。この場合、例えば、クエリーの遺伝子に対して公共データベースから配列情報を取得してsiRNAを選択してもよく、また、例えば、予め全遺伝子のsiRNAを計算して記憶しておき、クライアント端末からの要求（遺伝子の名称またはアクセッション番号など）に対して即座にsiRNAを選択し、選択したsiRNAをクライアント端末に返してもよい。

## 【0175】

また、塩基配列処理装置100は、標的遺伝子とは無関係の遺伝子に対する規定配列情報の特異性を確認してもよい。これにより、標的とする遺伝子のみの特異的にRNA干渉

を生じさせる規定配列情報を選出することができる。

【0176】

また、クライアント端末と塩基配列処理装置100とで構成される本システムには、例えば、Webページの利用者から、siRNAのRNA干渉効果（例えば、「効く」、「効かない」など）の結果をWeb上でフィードバックしてもらい、利用者からフィードバックされた実験例を塩基配列処理装置100に蓄積し、上述したRNA干渉に有効なsiRNAの配列規則性を改善するためのインターフェース機能を導入してもよい。

【0177】

また、塩基配列処理装置100は、規定配列情報からsiRNAのセンス鎖の塩基配列情報および当該センス鎖と相補的（complementary）なアンチセンス鎖の塩基配列情報を計算してもよい。具体的には、例えば、塩基配列処理装置100は、上述した各処理の結果、規定配列の両端に各2塩基のオーバーハング部を付加した23塩基の配列情報として「caccctgacccgcttcgtcatgg」を選択した場合、センス鎖の塩基配列情報「5'-CCCUGACCCGCUUCGUCAUGG-3'」およびアンチセンス鎖の塩基配列情報「5'-AUGACGAAGCGGGUCAGGGUG-3'」を計算する。これにより、ポリヌクレオチドを発注する時にセンス鎖、アンチセンス鎖を手作業でまとめる必要がなくなり、利便性が高まる。

【0178】

また、実施形態において説明した各処理のうち、自動的に行なわれるものとして説明した処理の全部または一部を手動的に行うこともでき、あるいは、手動的に行なわれるものとして説明した処理の全部または一部を公知の方法で自動的に行うこともできる。

【0179】

この他、上記文書中や図面中で示した処理手順、制御手順、具体的名称、各種の登録データや検索条件等のパラメータを含む情報、画面例、データベース構成については、特記する場合を除いて任意に変更することができる。

【0180】

また、塩基配列処理装置100に関して、図示の各構成要素は機能概念的なものであり、必ずしも物理的に図示の如く構成されていることを要しない。

【0181】

例えば、塩基配列処理装置100の各部または各装置が備える処理機能、特に制御部102にて行なわれる各処理機能については、その全部または任意の一部を、CPU（Central Processing Unit）および当該CPUにて解釈実行されるプログラムにて実現することができ、あるいは、ワイヤードロジックによるハードウェアとして実現することも可能である。なお、プログラムは、後述する記録媒体に記録されており、必要に応じて塩基配列処理装置100に機械的に読み取られる。

【0182】

すなわち、ROMまたはHDなどの記憶部106などには、OS（Operating System）と協働してCPUに命令を与え、各種処理を行うためのコンピュータプログラムが記録されている。このコンピュータプログラムは、RAM等にロードされることによって実行され、CPUと協働して制御部102を構成する。また、このコンピュータプログラムは、塩基配列処理装置100に対して任意のネットワーク300を介して接続されたアプリケーションプログラムサーバに記録されてもよく、必要に応じてその全部または一部をダウンロードすることも可能である。

【0183】

また、本発明にかかるプログラムを、コンピュータ読み取り可能な記録媒体に格納することもできる。ここで、この「記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、EPROM、EEPROM、CD-ROM、MO、DVD、フラッシュディスク等の任意の「可搬用の物理媒体」や、各種コンピュータシステムに内蔵されるROM、RAM、HD等の任意の「固定用の物理媒体」、あるいは、LAN、WAN、インターネットに代表されるネットワークを介してプログラムを送信する場合の通信回線や搬送波の



ように、短期にプログラムを保持する「通信媒体」を含むものとする。

【0184】

また、「プログラム」とは、任意の言語や記述方法にて記述されたデータ処理方法であり、ソースコードやバイナリコード等の形式を問わない。なお、「プログラム」は必ずしも単一的に構成されるものに限られず、複数のモジュールやライブラリとして分散構成されるものや、OS (Operating System) に代表される別個のプログラムと協働してその機能を達成するものをも含む。なお、実施の形態に示した各装置において記録媒体を読み取るための具体的な構成、読み取り手順、あるいは、読み取り後のインストール手順等については、周知の構成や手順を用いることができる。

【0185】

記憶部106に格納される各種のデータベース等(標的遺伝子塩基配列ファイル106a～標的遺伝子アノテーションデータベース106h)は、RAM、ROM等のメモリ装置、ハードディスク等の固定ディスク装置、フレキシブルディスク、光ディスク等のストレージ手段であり、各種処理やウェブサイト提供に用いる各種のプログラムやテーブルやファイルやデータベースやウェブページ用ファイル等を格納する。

また、塩基配列処理装置100は、既知のパーソナルコンピュータ、ワークステーション等の情報処理端末等の情報処理装置にプリンタやモニタやイメージスキャナ等の周辺装置を接続し、該情報処理装置に本発明の方法を実現させるソフトウェア(プログラム、データ等を含む)を実装することにより実現してもよい。

【0186】

さらに、塩基配列処理装置100等の分散・統合の具体的形態は明細書および図面に示すものに限られず、その全部または一部を、各種の負荷等に応じた任意の単位で、機能的または物理的に分散・統合して構成することができる(例えば、グリッド・コンピューティングなど)。例えば、各データベースを独立したデータベース装置として独立に構成してもよく、また、処理の一部をCGI (Common Gateway Interface) を用いて実現してもよい。

【0187】

また、ネットワーク300は、塩基配列処理装置100と外部システム200とを相互に接続する機能を有し、例えば、インターネットや、イントラネットや、LAN(有線/無線の双方を含む)や、VANや、パソコン通信網や、公衆電話網(アナログ/デジタルの双方を含む)や、専用回線網(アナログ/デジタルの双方を含む)や、CATV網や、IMT2000方式、GSM方式またはPDC/PDC-P方式等の携帯回線交換網/携帯パケット交換網や、無線呼出網や、Bluetooth等の局所無線網や、PHS網や、CS、BSまたはISDB等の衛星通信網等のうちいずれかを含んでもよい。すなわち、本システムは、有線・無線を問わず任意のネットワークを介して、各種データを送受信することができる。

【実施例】

【0188】

以下、実施例に基づき、本発明についてさらに詳細に説明する。なお、本発明は下記実施例に限定されるものではない。

【0189】

【実施例1】

<1>RNAi効果を測定するための遺伝子、発現ベクター

siRNAによるRNAi効果を測定するための標的遺伝子としてはホタル(Photinus pyralis, P. pyralis)のluciferase(luc)遺伝子(P. pyralis luc遺伝子: accession number: U47296)を用い、これを含む発現ベクターとしてはpGL3-Controlベクター(Promega社製)を用いた。P. pyralis luc遺伝子断片はこのベクター中でSV40のプロモーターとポリAシグナルで挟まれた形になっている。また内部コントロール遺伝子はウミシイタケ(Renilla reniformis, R



reniformis) の luc 遺伝子を用い、これを含む発現ベクターとして pRL-TK (Promega 社製) を用いた。

#### 【0190】

##### <2> 21塩基2本鎖RNA (siRNA) の合成

21塩基センス鎖と21塩基アンチセンス鎖RNA (第9図に示されている位置のもの; a~p) は、日立計測器サービス株式会社を通じてジェンセット株式会社に合成を委託した。

#### 【0191】

P. pyralis luc 遺伝子の発現を阻害するために用いた2本鎖RNAは、センス鎖とアンチセンス鎖とを会合させることで作製した。会合は、センス鎖RNAとアンチセンス鎖RNAを、10mM Tris-HCl (pH7.5), 20mM NaCl 反応液中で90℃、3分間加熱し、37℃で1時間インキュベートし、その後室温になるまで放置することで行った。2本鎖ポリヌクレオチドの形成は、TBE緩衝液中での2% アガロースゲルでの電気泳動で検定するが、上記条件では殆どすべての1本鎖ポリヌクレオチドが2本鎖ポリヌクレオチドに会合していた。

#### 【0192】

##### <3> 哺乳類細胞培養法

哺乳類培養細胞としては、ヒトのHeLa細胞とHEK293細胞、チャイニーズハムスターのCHO-K1細胞 (RIKEN Cell bank) を用いた。培地はDulbecco's modified Eagle's medium (Gibco BRL 社製) に非働化10%牛胎児血清 (Mitsubishi Kasei 社製) 及び抗生物質としてpenicillin (Meiji 社製) 10units/ml、streptomycin (Meiji 社製) 50μg/mlを添加したものをを用いた。37℃、5% CO<sub>2</sub>存在下で培養した。

#### 【0193】

##### <4> 哺乳類培養細胞への標的遺伝子、内部コントロール遺伝子、siRNAの培養細胞へのトランスフェクション

哺乳類細胞は0.2~0.3×10<sup>6</sup> cells/mlの濃度で24穴プレートにまき、1日後にCa-phosphate沈殿法 (細胞工学ハンドブック、黒木登志夫ら編、羊土社 (1992)) で1.0μg pGL3-Control DNA、0.5または1.0μg pRL-TK DNA、及び0.01、0.1、1、10、100nMの各siRNAを導入した。

#### 【0194】

##### <5> ショウジョウバエ細胞培養法

ショウジョウバエ培養細胞としてはS2細胞 (Schneider, I., et al., J. Embryol. Exp. Morph., 27, 353-365 (1972)) を用いた。培地はSchneider's Drosophila medium (Gibco BRL 社製) に非働化10%牛胎児血清 (Mitsubishi Kasei 社製) 及び抗生物質としてpenicillin (Meiji 社製) 10units/ml、streptomycin (Meiji 社製) 50μg/mlを添加したものをを用いた。25℃、5% CO<sub>2</sub>存在下で培養した。

#### 【0195】

##### <6> ショウジョウバエ培養細胞への標的遺伝子、内部コントロール遺伝子、siRNAの培養細胞へのトランスフェクション

このS2細胞は1.0×10<sup>6</sup> cells/mlの濃度で24穴プレートにまき、1日後にCa-phosphate沈殿法 (細胞工学ハンドブック、黒木登志夫ら編、羊土社 (1992)) で1.0μg pGL-3 Control DNA、0.1μg pRL-TK DNA、及び0.01、0.1、1、10、100nMの各siRNAを導入した。

#### 【0196】

## &lt;7&gt;RNAi効果の測定

siRNAをトランスフェクションした細胞は、トランスフェクション後20時間後に回収し、Dual-Luciferase Reporter Assay System (Promega社製)を用いて、2種類のルシフェラーゼ (P. pyralis luc及びR. reniformis luc) タンパクの発現量 (ルシフェラーゼ活性) を測定した。発光量の測定はLumat LB9507 luminometer (EG&G Berthold)を用いて行った。

【0197】

## &lt;8&gt;結果

ルシフェラーゼ活性を測定した結果を第10図に示す。また、ルシフェラーゼ活性と各塩基配列との対応を検討した結果を第11図に示す。

【0198】

第10図中、Bで示されるグラフはショウジョウバエでの結果であり、Cで示されるグラフはヒトでの結果を示す。第10図に示されるように、ショウジョウバエでは塩基数を21のRNAを作製することにより、殆どの配列においてルシフェラーゼ活性を抑制できた。他方、ヒトの場合については、単に塩基数を21とただけではルシフェラーゼ活性を抑制できる配列が得られにくいことが明らかとなった。

【0199】

そこで、a~pのRNAの塩基配列の規則性を分析した。配列の分析は、第11図に示すように2本鎖RNAの5箇所について塩基配列を分析した。第11図の表中最上段のaのsiRNAについてみると、ルシフェラーゼ相対活性 (RLA) は0.03、アンチセンス鎖について3'末端側から順に見ると、オーバーハング部 (OH) の塩基配列がUC、続く7塩基 (図11中3'-T) 中のG/C含有量 (グアニンまたはシトシンの含有量) は57%、さらに続く5塩基 (第11図中M) のG/C含有量は20であり、さらに続く7塩基 (第11図中5'-T) のG/C含有量は14%であり、5'末端はUであり、合計のG/C含有量は32%である。表中、RLAの値が低いほど、RLAの活性が低く、すなわちルシフェラーゼの発現が抑制されたことを示す。

【0200】

この結果から、RNA干渉を生じさせるポリヌクレオチドの塩基配列は、3'末端がアデニンまたはウラシルであり5'末端がグアニンまたはシトシンである確率が極めて高いことが明らかになった。また、3'末端側の7塩基についてはアデニンまたはウラシルが豊富な状態であることが明らかとなった。

【0201】

## [実施例2]

## 1. 標的発現ベクターpTRE Cの構築

下記のようにして標的発現ベクターを構築した。標的発現分子は、RNAiの標的となる配列 (以下「標的配列」という場合がある) を有するRNAを発現させる分子である。

【0202】

pCI-neo (GenBank AccessionNo.U47120、プロメガ社製) のCMVエンハンサー/プロモーター下流に標的mRNA配列を構築した (第25図)。すなわち、コザック配列 (Kozak)、ATG配列、標的となる23塩基対配列クロニングサイト (target) 及び組み換えのための制限酵素 (NheI、EcoRI、XhoI) 認識配列を有する下記の2本鎖オリゴマーを合成した。2本鎖オリゴマーは配列表の配列番号1に示す配列とその相補的な配列とからなる。合成した2本鎖オリゴマーをpCI-neoのNheI、XbaIサイトに挿入し、標的発現ベクターpTRE Cを構築した (第25図)。なお、イントロンは、pCI-neoに当初から組み込まれているβ-グロビン由来のイントロン部位を用いた。

【0203】

5'-gctagccaccatggaattcacgcgtctcgagtctaga-3' (配列番号1)

【0204】

第25図に示す pTREC は、プロモーターおよびエンハンサー (pro/enh)、と PCR プライマーに対応する領域 PAR (F) 1 および PAR (R) 1 を備える。PAR (F) 1 にはイントロン (Intron) が挟み込まれており、発現ベクターそのものが PCR の鋳型とならないように設計されている。pTREC は、RNA の転写後、真核生物の培養細胞などのスプライシングが行われる環境下において、イントロン部位が除去され、PAR (F) 1 が連結される。pTREC から形成された RNA は RT-PCR により増幅できる。イントロンは、pCIneo に当初から組み込まれている  $\beta$ -グロビン由来のイントロン部位を用いた。

#### 【0205】

pTREC には、コントロールとしてネオマイシン耐性遺伝子 (neo) が組み込まれており、ネオマイシン耐性遺伝子内の配列の一部に対応する PCR プライマーを作製し、ネオマイシン耐性遺伝子の一部について RT-PCR を行うことにより、ネオマイシン耐性遺伝子を内部標準コントロール (インターナルコントロール) として用いることができる。PAR (F) 2 および PAR (R) 2 は、ネオマイシン耐性遺伝子中の PCR プライマー対応領域を示す。なお、第25図の例には示していないが、PAR (F) 2 または PAR (R) 2 の少なくとも一方にもイントロンを挟み込んでおくこともできる。

#### 【0206】

##### 2. 標的 mRNA 検出用プライマーの効果

###### (1) 培養細胞へのトランスフェクション

24 穴プレートの 1 穴あたりに  $0.2 \sim 0.3 \times 10^6$  cells の HeLa 細胞をまき、1 日後に Lipofectamine 2000 (Invitrogen 社製) を用いて、マニュアルにしたがって  $0.5 \mu\text{g}$  の pTREC ベクターをトランスフェクトした。

#### 【0207】

###### (2) 細胞の回収および mRNA の定量

トランスフェクションの 1 日後、細胞を回収し、Trizol (Invitrogen 社製) を用いて全 RNA を抽出した。この RNA  $100 \text{ ng}$  を用いて、オリゴ (dT) をプライマーとし、SuperScript

II RT (Invitrogen 社製) により逆転写して cDNA 合成反応を行った。コントロールとして、逆転写酵素を加えないものを用意した。得られた cDNA の 320 分の 1 量を PCR テンプレートとし、SYBR

Green PCR Master Mix (Applied Biosystems 社製) を用いて、 $50 \mu\text{l}$  の反応系で定量的 PCR を行い、標的 mRNA (mRNA (T) と称する) および、内部コントロールとして pTREC 上のネオマイシン耐性遺伝子に由来する mRNA (mRNA (C) と称する) を定量した。定量的 PCR にはリアルタイムモニタリング装置 ABI

PRIZM7000 (Applied Biosystems 社製) を用い、mRNA (T) の定量にはプライマー対 T (配列表配列番号 2、3)、mRNA (C) の定量にはプライマー対 C (配列表配列番号 4、5) をそれぞれ使用した。

#### 【0208】

プライマー対 T:

aggcactgggcaggtgtc (配列番号 2)

tgctcgaagcattaaccctcacta (配列番号 3)

プライマー対 C

atcaggatgatctggacgaag (配列番号 4)

ctcttcagcaatatcacgggt (配列番号 5)

#### 【0209】

第26図および第27図に、PCR の結果を示す。第26図および第27図は、縦軸にそれぞれの PCR 産物を、横軸に PCR のサイクル回数を取り、グラフに表したものである。ネオマイシン耐性遺伝子の場合、逆転写酵素により cDNA を合成したもの (+RT) と、逆転写酵素を加えないコントロール (-RT) で、得られた PCR 産物量の増幅の差は小さかった (第26図)。これは、cDNA のみならず、細胞内に残っているベクタ

ーもPCRの鋳型となって増幅されたことを示す。他方、標的配列mRNAでは、逆転写酵素を添加した場合(+RT)と添加しない場合(-RT)とによる差が大きかった(第27図)。この結果は、プライマー対Tのうちの一方がイントロンをはさむ形でデザインされているため、イントロンが除去されたmRNAに由来するcDNAが効率的に増幅される一方、イントロンを有する残存ベクターが鋳型となりにくくなっていることを示している。

#### 【0210】

### 3. siRNAによる標的mRNAの発現抑制

#### (1) ターゲット発現ベクターへの評価配列のクローニング

ヒトビメンチン(VIM)遺伝子(RefSeqID: NM\_003380)コード領域の812-834、35-57に相当する配列を評価の対象とした。これらの配列およびEcoRI、XhoIの認識配列を有する下記配列表配列番号6および7の合成オリゴヌクレオチド(評価配列フラグメント)を作製した。

#### 【0211】

評価配列VIM35(VIMの35-57に相当)

5'-gaattcgcaggatgttcggcgcccgccgctcgag-3' (配列番号6)

評価配列VIM812(VIMの812-834に相当)

5'-gaattcacgtacgtcagcaatatgaaagtctcgag-3' (配列番号7)

#### 【0212】

得られた評価配列フラグメントの両端にあるEcoRI、XhoIサイトを利用し、pTRECのEcoRI、XhoIサイト間に、新たな標的配列としてクローニングし、pTREC-VIM35、pTREC-VIM812を構築した。

#### 【0213】

#### (2) siRNAの作製

評価配列VIM35(配列表配列番号8、第28図)、評価配列VIM812(配列番号9、第29図)、及びコントロール配列(siControl、配列番号10、第30図)にそれぞれ相当するsiRNAフラグメントを合成し、アニーリングした。下記siRNAの各配列においては、3'末端にオーバーハング部を設けてある。

#### 【0214】

siVIM35 5'-aggauugcggcgcccgccg-3' (配列番号8)

siVIM812 5'-guacgucagcaauaugaaagu-3' (配列番号9)

#### 【0215】

コントロールとして、ルシフェラーゼ遺伝子に対するsiRNAを用いた。

#### 【0216】

siControl 5'-cauucuaucgcuggaagaug-3' (配列番号10)

#### 【0217】

#### (3) 培養細胞へのトランスフェクション

24穴プレートの1穴あたりに $0.2 \sim 0.3 \times 10^6$  cellsのHeLa細胞をまき、1日後にLipofectamine 2000(Invitrogen社製)を用いて、マニュアルにしたがって $0.5 \mu\text{g}$ のpTREC-VIM35またはpTREC-VIM812と、それぞれのVIM由来配列に相当する $100 \text{ nM}$ のsiRNA(siVIM35、siVIM812)を同時にトランスフェクトした。コントロール細胞には、 $0.5 \mu\text{g}$ のpTREC-VIM35またはpTREC-VIM812と $100 \text{ nM}$ のルシフェラーゼ遺伝子に対するsiRNA(siControl)を同時にトランスフェクトした。

#### 【0218】

#### (4) 細胞の回収およびmRNAの定量

トランスフェクションの1日後、細胞を回収し、Trizol(Invitrogen)を用いて全RNAを抽出した。このRNA $100 \text{ ng}$ を用いて、オリゴ(dT)をプライマーとし、SuperScript

II RT(Invitrogen社製)により逆転写してcDNAを得た。得られたcDNAの $320$

分の1量をPCRテンプレートとし、SYBR Green PCR Master Mix(Applied Biosystems社製)を用いて、50  $\mu$ lの反応系で定量的PCRを行い、評価しようとするVIM由来の配列を含むmRNA (mRNA (T) と称する) および内部コントロールとして、

pTREC上のネオマイシン耐性遺伝子に由来するmRNA (mRNA (C) と称する) を定量した。

#### 【0219】

定量的PCRにはリアルタイムモニタリング装置ABI PRISM7000(Applied Biosystems社製)を用い、mRNA (T) の定量にはプライマー対T (配列表配列番号2および3)、mRNA (C) の定量にはプライマー対C (配列表配列番号4および5) をそれぞれ使用した。得られたそれぞれのmRNAの値の比(T/C)を縦軸(標的mRNA相対量(%))としてグラフに表した(第31図)。

#### 【0220】

コントロール細胞の場合、ルシフェラーゼ遺伝子に対するsiRNAは標的mRNAに効果を及ぼさないため、T/C比はほぼ1になった。VIM812 siRNAでは、T/C比が著しく下がった。これは、VIM812 siRNAが、相当する配列を有するmRNAを切断したためであり、VIM812 siRNAがRNAi効果を有することが示された。一方、VIM35 siRNAの場合、T/C比はコントロールとほぼ同じ値であったことから、VIM35の配列にはRNAi効果がほとんどないことが示された。

#### 【0221】

##### [実施例3]

##### 1. siRNAによる内在性ビメンチンの発現抑制

##### (1) 培養細胞へのトランスフェクション

24穴プレートの1穴あたりに0.2~0.3  $\times 10^6$  cellsのHeLa細胞をまき、1日後にLipofectamine 2000 (Invitrogen社製)を用いて、マニュアルにしたがって、VIMに対するsiRNA (siVIM35またはsiVIM812) またはコントロールsiRNA (siControl) 100 nM、トランスフェクション効率のコントロールとして0.5  $\mu$ gのpEGFP (Clontech社製)を同時にトランスフェクトした。pEGFPはEGFPが組み込まれている。

#### 【0222】

##### (2) 内在性ビメンチンmRNAの測定

トランスフェクションの3日後、細胞を回収し、Trizol(Invitrogen社製)を用いて全RNAを抽出した。このRNA 100 ngを用いて、オリゴ(dT)をプライマーとし、SuperScript

II RT(Invitrogen社製)により逆転写してcDNA合成反応を行なった。得られたcDNA産物を鋳型として、ビメンチンに対するプライマーVIM-F3-84、VIM-R3-274 (配列番号11、12)を用いてPCRを行った。

#### 【0223】

VIM-F3-84; gagctacgtgactacgtcca (配列番号11)

VIM-R3-274; gttcttgaactcgggtgtgat (配列番号12)

#### 【0224】

また、コントロールとして、 $\beta$ -アクチンに対するプライマーACTB-F2-481、ACTB-R2-664 (配列番号13、14)を用いてPCRを行い、各サンプル間の $\beta$ -アクチンの定量値を合わせたうえで、ビメンチンの発現量を評価した。

#### 【0225】

ACTB-F2-481; cacactgtgcccatctacga (配列番号13)

ACTB-R2-664; gccatctcttgctcgaagtc (配列番号14)

#### 【0226】

結果を第32図に示す。第32図では、siControl (すなわち、標的とは無関

係な配列)を入れた場合を100%として比較し、VIMに対するsiRNAを入れた場合に、VIMのmRNAがどの程度減少したかを表している。siVIM-812は効果的にVIM mRNAを抑制できたのに対して、siVIM-35を用いた場合はほとんどRNAi効果を示さなかった。

【0227】

### (3) 細胞の抗体染色

トランスフェクションの3日後、細胞を3.7%のホルムアルデヒドにより固定し、定法によりブロッキングを行った。その後、ウサギ抗ビメンチン抗体( $\alpha$ -VIM)または内部コントロールとしてウサギ抗Yes抗体( $\alpha$ -Yes)を加え、室温で反応させた。その後、細胞表面をPBS (Phosphate Buffered Saline; リン酸緩衝生理食塩水)で洗浄し、二次抗体として蛍光標識抗ウサギIgG抗体を加え、室温で反応させた。細胞表面をPBSで洗浄の後、蛍光顕微鏡による観察を行った。

【0228】

蛍光顕微鏡観察の結果を第33図に示す。第33図中、9つの各枠内において白く現れている部分が蛍光部分である。EGFPおよびYesについては、何れの細胞でも同程度の発現が確認された。siControl、siVIM35を導入した細胞では、ビメンチンの抗体染色による蛍光が観察され、内在性ビメンチンの存在が確認された。一方、siVIM812を導入した細胞では、siControl、siVIM35を導入した細胞に比べて蛍光が著しく弱かった。この結果は、siVIM812により内在性のビメンチンmRNAが干渉を受けた結果、ビメンチンの蛋白質の発現量が減少したことを示すものであり、siVIM812は内在性ビメンチンmRNAにもRNAi効果を有することが明らかになった。

【0229】

本発明のアッセイ系で得られた結果[実施例2]は、実際に内在性遺伝子に対してそれぞれのsiRNAを用いた結果[実施例3]とよく一致していたため、このアッセイ系は任意のsiRNAのRNAi活性を評価する方法として有効であることが示された。

【0230】

### [実施例4]

上記所定の規則(a)~(d)に基づき塩基配列を設計した。塩基配列の設計は、上記siRNA配列設計プログラムを実行する塩基配列処理装置によって行った。塩基配列は、RNAi活性を有すると予測される配列15種(配列番号15~29)と、RNAi活性を有しないと予測される配列5種(配列番号30~34)を用意した。

【0231】

設計された配列に基づき標的配列および評価対象のsiRNAを調製する以外は、上記実施例1と同様にして、ルシフェラーゼ活性を測定することにより、RNAi活性の評価を行った。結果を第34図に示す。ルシフェラーゼ相対活性値が低いものが効いている状態、すなわちRNAi活性を備えたsiRNAである。上記プログラムによりRNAi活性を有すると予測されたsiRNAは、全てルシフェラーゼの発現を効果的に抑制した。

【0232】

[RNAi活性を示した配列; 規定配列部分、オーバーハング部を含まない]

- 5, gacgccaaaacataaaga (配列番号15)
- 184, gttggcagaagctatgaaa (配列番号16)
- 272, gtgttgggcgcgttatatta (配列番号17)
- 309, ccgccaacgacatttataa (配列番号18)
- 428, ccaatcatccaaaaatta (配列番号19)
- 515, cctcccggttttaatgaat (配列番号20)
- 658, gcatgccagagatcctatt (配列番号21)
- 695, ccgatactgcgattttaa (配列番号22)
- 734, ggttttggaatgtttacta (配列番号23)

774, gatttcgagtcgtcttaat (配列番号 2 4)  
891, gcactctgattgacaaata (配列番号 2 5)  
904, caaatacgatttatcta (配列番号 2 6)  
1186, gattatgtccggttatgta (配列番号 2 7)  
1306, ccgctgaagtctctgatt (配列番号 2 8)  
1586, ctcgacgcaagaaaaatca (配列番号 2 9)

## 【0 2 3 3】

[RNA i 活性を示さなかった配列; 規定配列部分、オーバーハング部を含まない]

14, aacataaagaaaggcccg (配列番号 3 0)  
265, tatgccggtgttggcgcg (配列番号 3 1)  
295, agttgcagttgcgcccg (配列番号 3 2)  
411, acgtgcaaaaaagctccc (配列番号 3 3)  
1044, ttctgattacacccgaggg (配列番号 3 4)

## 【0 2 3 4】

## [実施例 5]

SARS ウイルスに対する siRNA を設計し、それらの RNA i 活性を調べた。標的配列および評価対象の配列をそれぞれ変更する以外は、RNA i 活性は上記実施例 2 と同様のアッセイを行うことにより評価した。

## 【0 2 3 5】

siRNA は、SARS ウイルスのゲノムから、3CL-PRO、RdRp、Spike glycoprotein、Small envelope E protein、Membrane glycoprotein M、Nucleocapsid protein、s2m motif の各所について、上記 siRNA 配列設計プログラムを用い所定の規則性に合致するものを設計した。

第 3 5 図アッセイの結果、規則性に合致するように設計した 11 個の siRNA は、それぞれ対応する siRNA の配列をターゲットとして組み込んだ RNA を効果的に抑制した。siControl (SARS とは関係のない配列) を入れた場合を 100% として、SARS の各 siRNA を入れた場合の相対標的 mRNA 量を示す。各 siRNA を入れた場合、標的の RNA が 10% 程度あるいはそれ以下まで減少しており、RNA i 活性があることが確認された。

## 【0 2 3 6】

[設計した siRNA の配列 (規定配列部分、オーバーハング部を含まない)]

siControl; gggcgcggtcggtaaagtt (配列番号 3 5)  
3CL-PRO; SARS-10754; ggaattgccgtcttagata (配列番号 3 6)  
3CL-PRO; SARS-10810; gaatggtcggtactatcctt (配列番号 3 7)  
RdRp; SARS-14841; ccaagtaatcgtaacaat (配列番号 3 8)  
Spike glycoprotein; SARS-23341; gcttggcgcatatattcta (配列番号 3 9)  
Spike glycoprotein; SARS-24375; cctttcgcgacttgataaa (配列番号 4 0)  
Small envelope E protein; SARS-26233; gtgcgtactgctgcaatat (配列番号 4 1)  
Small envelope E protein; SARS-26288; ctactcgcggtgttaaaaat (配列番号 4 2)  
Membrane glycoprotein M; SARS-26399; gcagacaacggtactatta (配列番号 4 3)  
Membrane glycoprotein M; SARS-27024; ccggtagcaacgacaatat (配列番号 4 4)  
Nucleocapsid protein; SARS-28685; cgtagtcgcggttaattcaa (配列番号 4 5)  
s2m motif; SARS-29606; gatcgagggtacagtgaat (配列番号 4 6)

## 【0 2 3 7】

## [実施例 6]

上記「< 5 > siRNA 配列設計プログラム」および「< 7 > siRNA 配列設計プログラムを実行する塩基配列処理装置等」に従って、下記 siRNA を設計した。なお、プログラム実行の際の設定条件は、下記の通りである。

## 【0 2 3 8】

(設定条件)

- (a) 3'末端の塩基が、アデニン、チミンまたはウラシルである。
- (b) 5'末端の塩基が、グアニンまたはシトシンである。
- (c) 3'末端の7塩基の配列において、4塩基以上が、アデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上である。
- (d) 塩基数が、19である。
- (e) 10塩基以上グアニンまたはシトシンが連続する配列を含まない。
- (f) 目的生物の全遺伝子配列のうち、前記標的遺伝子以外の他の遺伝子の塩基配列中に、当該規定配列とのミスマッチ数が2塩基以下の類似配列が含まれない。

**【0239】**

設計された siRNA の配列を配列表の配列番号 47～配列番号 817081 に示す。なお、配列表の配列番号 47～配列番号 817081 に記載された、各 siRNA の目的生物の生物名は、配列表の〈213〉に記載されている。また、RNAi 対象遺伝子の遺伝子名、対象遺伝子の accession、および、対象遺伝子の塩基配列において規定配列に該当する部分については、配列表の〈223〉(Other information) に記載されている。なお、ここでの遺伝子名および accession 情報は、NCBI の「RefSeq」(HYPERLINK "<http://www.ncbi.nlm.nih.gov/>" <http://www.ncbi.nlm.nih.gov/>) に対応するものであり、RefSeq にアクセスすることによって、各遺伝子の情報(遺伝子の配列、機能等)を取得することができる。

**【0240】**

配列番号 47 に記載の siRNA を例にして説明すると、目的生物は Homo sapiens であり、対象遺伝子の遺伝子名は、ATBF1 であり、対象遺伝子の accession は、NM\_006885.2 であり、規定配列に該当する部分は、NM\_006885.2 の塩基配列のうち塩基番号 908-926 の 19 塩基である。RefSeq にアクセスすると、当該対象遺伝子は、AT-binding transcription factor 1 に関連する遺伝子であることが分かる。

**【0241】****[実施例 7]**

siRNA とのミスマッチ数が小さい配列を含む他の遺伝子への影響を調べるため、実施例 5 と同様の手法で、ホタルルシフェラーゼに対する siRNA を設計し、当該 siRNA とミスマッチ数が小さい類似配列に対する RNAi 効果を調べた。

**【0242】**

[設計した siRNA の配列(規定配列部分、塩基数 2 のオーバーハング部を含む)]  
3-36 gccattctatccgctggaagatg (配列番号 817082)

**【0243】**

[設計した siRNA の類似配列(大文字で表記した塩基がミスマッチ部位)]

- 3-36.R1 gccattctatccgcGggCGgatg (配列番号 817083)
- 3-36.R2 gccattctatccgcCggGGgatg (配列番号 817084)
- 3-36.R3 gccattctatccgcGggaCgatg (配列番号 817085)
- 3-36.R4 gccattctatccgctggCGgatg (配列番号 817086)
- 3-36.R5 gccattctatccgctggaGgatg (配列番号 817087)
- 3-36.R6 gccattctatccgctgTaaTatg (配列番号 817088)
- 3-36.R7 gccattctatccgctAAaagatg (配列番号 817089)
- 3-36.R8 gccattctatccgctATaaAatg (配列番号 817090)
- 3-36.L1 gccGGCcCGtccgctggaagatg (配列番号 817091)
- 3-36.L2 gccCGtCtCGtccgctggaagatg (配列番号 817092)
- 3-36.L3 gccGtCctGtccgctggaagatg (配列番号 817093)
- 3-36.L4 gccACCcGatccgctggaagatg (配列番号 817094)
- 3-36.L5 gccattAtatccgctggaagatg (配列番号 817095)
- 3-36.01A gcAattctatccgctggaagatg (配列番号 817096)
- 3-36.01G gcGattctatccgctggaagatg (配列番号 817097)



3-36.01U gcTattctatccgctggaagatg (配列番号 817098)  
3-36.19G gccattctatccgctggaagGtg (配列番号 817099)  
3-36.19C gccattctatccgctggaagCtg (配列番号 817100)  
3-36.19U gccattctatccgctggaagTtg (配列番号 817101)

【0244】

第36図に示すアッセイの結果、塩基数19の塩基配列を設計する場合、標的遺伝子以外の他の遺伝子がミスマッチ数2塩基以下の類似配列を含む場合、当該類似配列部分がRNA干渉の標的となりうる可能性が高いことが確認された。

【産業上の利用可能性】

【0245】

以上のように、本発明にかかるポリヌクレオチドは、標的遺伝子に対して高いRNA干渉効果を有するとともに、標的遺伝子と関係のない遺伝子についてRNA干渉を生じる危険性が非常に小さいため、発現を抑制しようとする標的遺伝子だけに特異的にRNA干渉を生じさせることができる。したがって、本発明は、RNA干渉を利用する試験、治療方法等に好適に利用することができ、哺乳類などの高等動物、特にヒトを対象とするRNA干渉を行う際に特に有効性を発揮するものである。

【図面の簡単な説明】

【0246】

【図1】 ヒトとマウスとで共通配列である s-iRNA の設計を示す図である。

【図2】 RNAi 効果を示す s-iRNA の規則性を示す図である。

【図3】 ヒトFBP1およびマウスFbp1の塩基配列中の規定配列を有する共通部位(太字部分)を示す図である。

【図4】 ヒトFBP1およびマウスFbp1に共通の規定配列を列挙した図である。

【図5】 ヒトFBP1およびマウスFbp1に共通の規定配列にスコアを付した図である。

【図6】 標的以外の遺伝子をノックアウトしないよう、規定配列のうちの1つをBLAST検索した結果を示す図である。

【図7】 標的以外の遺伝子をノックアウトしないよう、規定配列のうちの1つをBLAST検索した結果を示す図である。

【図8】 プログラムの出力結果を示す図である。

【図9】 RNA断片の設計(a~p)を示す図である。

【図10】 a~pのs-iRNAがRNAi効果を示すか試験した結果を示す図である。「B」はショウジョウバエ培養細胞における結果を、「C」はヒト培養細胞における結果を示す図である。

【図11】 a~pのs-iRNAの配列の特徴を分析した結果を示す図である。

【図12】 本発明の基本原理を示す原理構成図である。

【図13】 本発明が適用される本システムの塩基配列処理装置100の構成の一例を示すブロック図である。

【図14】 標的遺伝子塩基配列ファイル106aに格納される情報の一例を示す図である。

【図15】 部分塩基配列ファイル106bに格納される情報の一例を示す図である。

【図16】 判定結果ファイル106cに格納される情報の一例を示す図である。

【図17】 規定配列ファイル106dに格納される情報の一例を示す図である。

【図18】 参照配列データベース106eに格納される情報の一例を示す図である。

【図19】 同一類似度ファイル106fに格納される情報の一例を示す図である。

【図20】 評価結果ファイル106gに格納される情報の一例を示す図である。

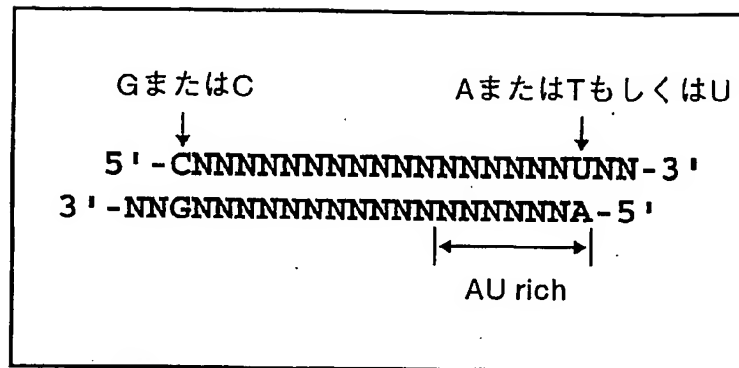
【図21】 本発明が適用される本システムの部分塩基配列作成部102aの構成の一例を示すブロック図である。

【図22】 本発明が適用される本システムの無関係遺伝子標的評価部102hの構成の一例を示すブロック図である。

- 【図23】本実施形態における本システムのメイン処理の一例を示すフローチャートである。
- 【図24】本実施形態における本システムの無関係遺伝子標的評価処理の一例を示すフローチャートである。
- 【図25】標的発現ベクター pTREC の構造を示す図である。
- 【図26】実施例2の2. (2) におけるプライマーの一方がイントロンを挟む形でデザインされていない場合のPCRの結果を示す図である。
- 【図27】実施例2の2. (2) におけるプライマーの一方がイントロンを挟む形でデザインされている場合のPCRの結果を示す図である。
- 【図28】siRNA; siVIM35の配列および構造を示す図である。
- 【図29】siRNA; siVIM812の配列および構造を示す図である。
- 【図30】、siRNA; siControlの配列および構造を示す図である。
- 【図31】siVIM812およびsiVIM35のRNAi活性をアッセイした結果を示す図である。
- 【図32】siControl、siVIM812およびsiVIM35のビメンチンに対するRNAi活性を示す図である。
- 【図33】抗体染色の結果を示す図である。
- 【図34】プログラムにより設計された siRNA のルシフェラーゼ遺伝子に対する RNA i 活性の測定結果を示す図である。
- 【図35】プログラムにより設計された siRNA の SARS ウイルスが有する配列に対する RNA i 活性の測定結果を示す図である。
- 【図36】プログラムにより設計された siRNA とのミスマッチ数が小さい配列を含む他の遺伝子に対する RNA i 活性の測定結果を示す図である。
- 【0247】



【圖 2】



【図 3】

		NM_000507: <i>Homo sapiens</i> fructose-1,6-bisphosphatase 1 ( <i>FBP1</i> )		NM_019395: <i>Mus musculus</i> fructose bisphosphatase 1 ( <i>Fbp1</i> )			
NM_000507	1	ATGGCTGACCAGGCGCCCTTCGACACGGACGTCAACACCCTGACCCGCTTCGTCATGGAG	60	NM_019395	1	ATGGCGAACCATGCGCCCTTCGAAACGGATATCAGCACCCCTGACCCGCTTCGTCATGGAG	60
		*****					
NM_000507	61	GAGGGCAGGAAGGCCCGCGGCACGGGCGAGTTGACCCAGCTGCTCAACTCGCTCTGCACA	120	NM_019395	61	CAGGGCAGGAAGGCTCAGGGCAGGGGAGTTGACCCAGCTGCTGAATTCGCTCTGCACC	120
		*****					
NM_000507	121	GCAGTCAAAGCCATCTCTTCGGCGGTGCGCAAGGCGGGCATCGCGCACCTCTATGGCATT	180	NM_019395	121	GCGATCAAAGCCATCTCGTCTGCGGTGCGCCAGGCGGGCATCGCACAGCTCTATGGTATC	180
		**					
		▼ intron					
NM_000507	181	GCTGGTCTTACCAACGTGACAGGTGATCAAGTTAAGAAGCTGGACGCTCTCTCCAACGAC	240	NM_019395	181	GCTGGCTCAACCAATGTGACTGGGGATCAAGTAAAGAAGCTGGACATACTTTCCAATGAC	240
		*****					
NM_000507	241	CTGGTTATGAACATGTTAAAGTCATCCTTTGCCACGTGTGTTCTCGTGTGAGAAGAAGAT	300	NM_019395	241	CTGGTGATCAATATGCTGAAGTCGTCCTACGCTACCTGTGTTCTTGTGCTGAAGAAAAC	300
		*****					
		▼ intron					
NM_000507	301	AAACACGCCATCATAGTGGAAACCGAGAAAAGGGGTAAATATGTGGTCTGTTTGTATCCC	360	NM_019395	301	ACAAATGCCATCATATCGAACCTGAGAAGAGGGGGCAAAATATGTTGTCTGTTTTCATCCC	360
		* * *					
		▲					
NM_000507	361	CTTGATGGATCTTCCAACATCGATTGCGCTTGTGTCCGTTGGAACCATTTTGGCATCTAT	420	NM_019395	361	CTTGATGGCTCATCCAACATTGACTGCGCTTGTGTCCATCGGAACCATTTTGGCATTAC	420
		*****					
		▼ intron					
NM_000507	421	AGAAAGAAATCAACTGATGAGCCTTCTGAGAAGGATGCTCTGCAACCAGGCCGGAACCTG	480	NM_019395	421	AGAAAGAAAGTACTGATGAGCCTTCTGAGAAGGATGCTCTGCAGCCCGGCGGGACCTG	480
		*****					
		▲					
NM_000507	481	GTGGCAGCCGGCTACGCACTGTATGGCAGTGCCACCATGCTGGTCCCTTGCCATGGACTGT	540	NM_019395	481	GTGGCAGCCGGGTATGCGCTCTATGGCAGTGCCACCATGTTGGTCCCTTGCCATGGATTGT	540
		*****					
		▼ intron					
NM_000507	541	GGGGTCAACTGCTTCATGCTGGACCCCGCCATCGGGGAGTTCAATTTGGTGGACAAGGAT	600	NM_019395	541	GGTGTCAACTGCTTCATGCTGGACCCGCTCCATTGGAGAATTCATTATGGTGGACAGGGAC	600
		**					
		▲					
NM_000507	601	GTGAAGATAAAAAAGAAAGTAAATCTACAGCCTTAACGAGGGCTACGCCAGGGACTTT	660	NM_019395	601	GTGAAGATGAAGAAGAAAGTAAATCTACAGCCTTAATGAGGGTTATGCCAAGGACTTT	660
		*****					
		▼ intron					
NM_000507	661	GACCTGCGCTCACTGAGTACATCCAGAGGAAGAAGTTCCTCCAGATAATTCAGCTCCT	720	NM_019395	661	GACCTGCCATCAATGAGTATCTCCAGAGGAAAAGTTCCTCCGGATGGTTTACGCCCC	720
		*****					
		▲					
NM_000507	721	TATGGGCCCCGATATGTGGGCTCCATGGTGGCTGATGTTTCATCGCACTCTGGTCTACGGA	780	NM_019395	721	TATGGTGCCCGGTATGTGGGCTCCATGGTGGCTGATATTCACCGCACTCTGGTATATGGA	780
		*****					
		▼ intron					
NM_000507	781	GGGATATTTCTGTATCCCGCTAACAAGAAGAGCCCCAATGGAAAGCTGAGACTGCTGTAC	840	NM_019395	781	GGGATCTTTTATACCCCGCCAAACAAGAAAAGCCCAAGTGGAAAGCTGCGGCTGCTGTAT	840
		*****					
		▲					
NM_000507	841	GAATGCAACCCCATGGCCTACGTCATGGAGAAGGCTGGGGGAATGGCCACCACTGGGAAG	900	NM_019395	841	GAGTGCAACCCCATAGCTTATGTCTATGGAGAAGGCGGTGGGCTCGCCACCACGGGGAC	900
		**					
NM_000507	901	GAGGCCGTGTAGACGTCTATCCACAGACATTACCCAGAGGGCGCCGGTGTCTTGGGA	960	NM_019395	901	AAAGATATATTAGACATCGTTCCACCGAGATCCACCAGAAGGCACCACTCGTCATGGGG	960
		* * *					
NM_000507	961	TCCCCGACGACGTGCTCGAGTCTCTGAAGGTGTATGAGAAGCACTCTGCCCACTGA	1017	NM_019395	961	TCCTCTGAAGATGTGCAGGAGTCTCTGGAGATCTACAGGAAGCAAAAGCCAAAGTGA	1017
		***					

【図 4】

```
NM_000507: 36 caccctgacccgcttcgtcatgg
NM_000507: 37 accctgacccgcttcgtcatgga
NM_000507: 38 ccctgacccgcttcgtcatggag
NM_000507:433 actgatgagccttctgagaagga
NM_000507:434 ctgatgagccttctgagaaggat
NM_000507:435 tgatgagccttctgagaaggatg
NM_000507:436 gatgagccttctgagaaggatgc
NM_000507:437 atgagccttctgagaaggatgct
NM_000507:438 tgagccttctgagaaggatgctc
NM_000507:439 gagccttctgagaaggatgctct
NM_000507:440 agccttctgagaaggatgctctg
NM_000507:441 gccttctgagaaggatgctctgc
NM_000507:442 ccttctgagaaggatgctctgca
NM_000507:544 gtcaactgcttcatgctggaccc
NM_000507:545 tcaactgcttcatgctggacccg
```

【図 5】

```
NM_000507: 36 caccctgacccgcttcgtcatgg 1 1 4 4
NM_000507: 37 accctgacccgcttcgtcatgga 0 1 3 0
NM_000507: 38 ccctgacccgcttcgtcatggag 0 1 3 0
NM_000507:433 actgatgagccttctgagaagga 0 0 4 0
NM_000507:434 ctgatgagccttctgagaaggat 0 1 3 0
NM_000507:435 tgatgagccttctgagaaggatg 1 0 4 0
NM_000507:436 gatgagccttctgagaaggatgc 1 0 4 0
NM_000507:437 atgagccttctgagaaggatgct 0 1 4 0
NM_000507:438 tgagccttctgagaaggatgctc 0 0 3 0
NM_000507:439 gagccttctgagaaggatgctct 1 1 3 0
NM_000507:440 agccttctgagaaggatgctctg 0 1 3 0
NM_000507:441 gccttctgagaaggatgctctgc 1 1 4 4
NM_000507:442 ccttctgagaaggatgctctgca 0 0 3 0
NM_000507:544 gtcaactgcttcatgctggaccc 0 1 2 0
NM_000507:545 tcaactgcttcatgctggacccg 0 0 2 0
```

【図6】

## Sequences producing significant alignments:

		Score (bits)	E Value
ref NM_019395.1	Mus musculus fructose biphosphatase 1 (Fbp1), ...	46	9e-06
ref NM_000507.2	Homo sapiens fructose-1,6-bisphosphatase 1 (FBP)...	46	9e-06
ref NM_015820.1	Mus musculus heparan sulfate 6-O-sulfotransfera...	30	0.52
ref NM_003837.1	Homo sapiens fructose-1,6-bisphosphatase 2 (FBP)...	30	0.52
ref NM_145960.1	Mus musculus hypothetical protein MGC19099 (MGC...	28	2.0
ref NM_010884.1	Mus musculus N-myc downstream regulated 1 (Ndr1)...	28	2.0
ref NM_025405.1	Mus musculus RIKEN cDNA 1110033J19 gene (111003...	26	8.1
ref NM_029098.1	Mus musculus RIKEN cDNA 1110013E13 gene (111001...	26	8.1
ref NM_012258.2	Homo sapiens hairy/enhancer-of-split related wi...	26	8.1
ref NM_018113.1	Homo sapiens lipocalin-interacting membrane rec...	26	8.1
ref NM_014780.1	Homo sapiens KIAA0076 gene product (KIAA0076), ...	26	8.1
ref NM_012218.1	Homo sapiens interleukin enhancer binding facto...	26	8.1
ref NM_004516.1	Homo sapiens interleukin enhancer binding facto...	26	8.1
ref NM_000875.2	Homo sapiens insulin-like growth factor 1 recep...	26	8.1
ref NM_001188.1	Homo sapiens BCL2-antagonist/killer 1 (BAK1), mRNA	26	8.1

【図7】

## Sequences producing significant alignments:

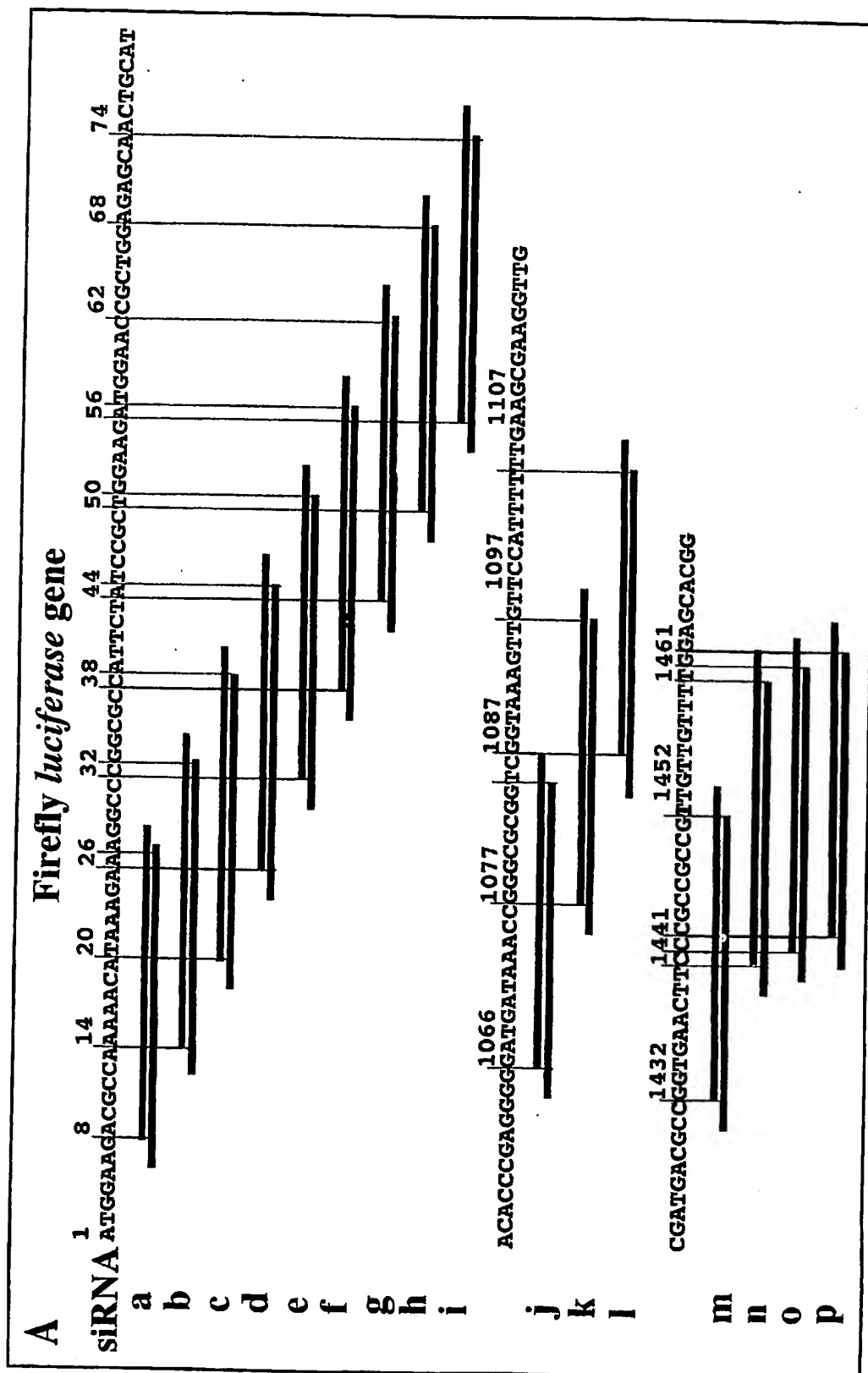
		Score (bits)	E Value
ref NM_019395.1	Mus musculus fructose biphosphatase 1 (Fbp1), ...	46	9e-06
ref NM_000507.2	Homo sapiens fructose-1,6-bisphosphatase 1 (FBP...)	46	9e-06
ref NM_007994.1	Mus musculus fructose biphosphatase 2 (Fbp2), ...	36	0.008
ref NM_139045.1	Homo sapiens SWI/SNF related, matrix associated...	32	0.13
ref NM_003070.2	Homo sapiens SWI/SNF related, matrix associated...	32	0.13
ref NM_022026.1	Mus musculus aquaporin 9 (Aqp9), mRNA	30	0.52
ref NM_011834.1	Mus musculus kynurenine aminotransferase II (Ka...	30	0.52
ref NM_008632.1	Mus musculus microtubule-associated protein 2 (...)	30	0.52
ref NM_023266.1	Mus musculus RIKEN CDNA 1200003I07 gene (120000...	30	0.52
ref NM_144499.1	Homo sapiens guanine nucleotide binding protein...	30	0.52
ref NM_000172.2	Homo sapiens guanine nucleotide binding protein...	30	0.52
ref NM_015125.2	Homo sapiens capicua homolog (Drosophila) (CIC)...	30	0.52
ref NM_146249.1	Mus musculus hypothetical protein MGC18735 (MGC...	28	2.0
ref NM_027017.1	Mus musculus RIKEN CDNA 3300002I08 gene (330000...	28	2.0
ref NM_145490.1	Mus musculus similar to zinc finger protein 97 ...	28	2.0
ref NM_145528.1	Mus musculus similar to Hypothetical protein KI...	28	2.0
ref NM_144546.1	Mus musculus expressed sequence AL024077 (AL024...	28	2.0
ref NM_144852.1	Mus musculus hypothetical protein MGC27672 (MGC...	28	2.0
ref NM_080467.1	Mus musculus ATPase, H+ transporting, lysosomal...	28	2.0
ref NM_029813.1	Mus musculus RIKEN CDNA 2210418010 gene (221041...	28	2.0
ref NM_029583.1	Mus musculus RIKEN CDNA 2810408B13 gene (281040...	28	2.0
ref NM_001231.2	Homo sapiens calsequestrin 1 (fast-twitch, skel...	28	2.0
ref NM_020824.1	Homo sapiens Rho-GTPase activating protein 10 (...)	28	2.0
...			



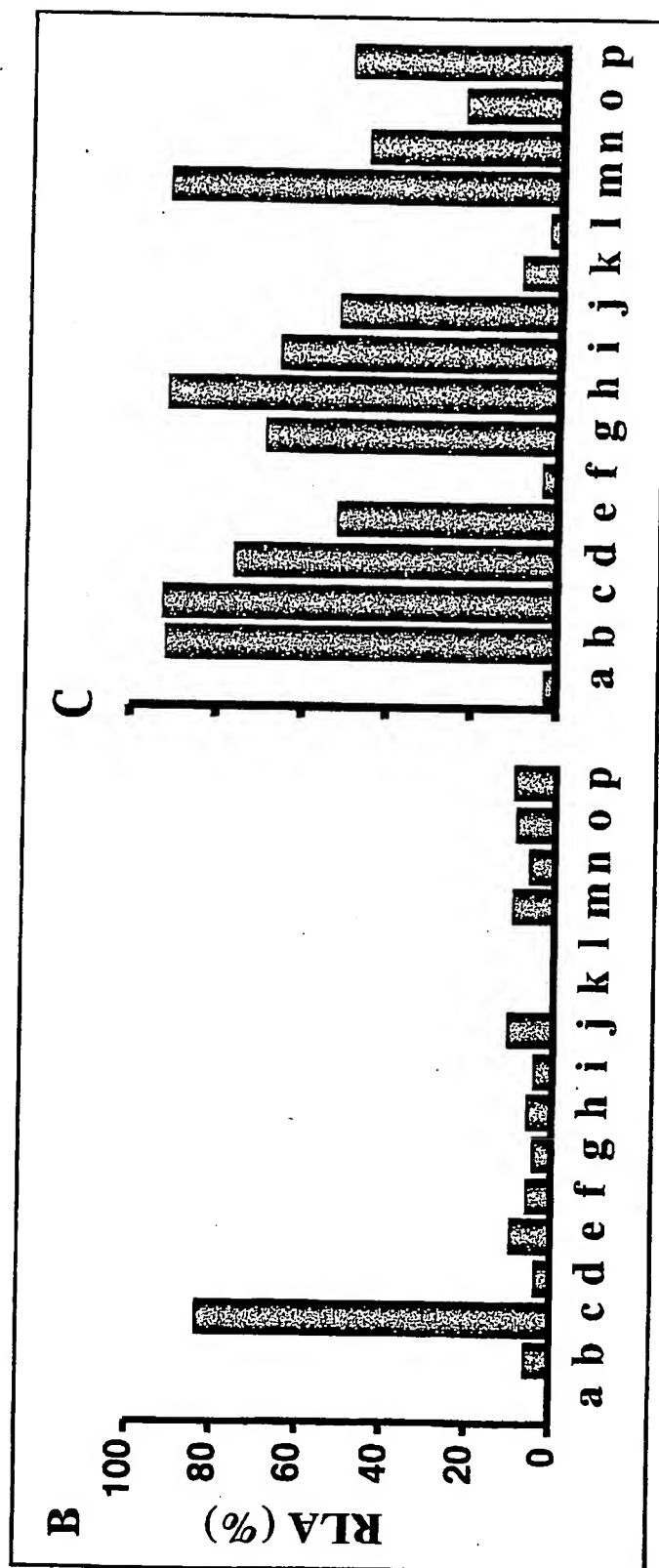
【図8】

NM_000507: 36	caccctgacccgcttcgtcatgg	1	1	4	4	59
NM_000507: 37	accctgacccgcttcgtcatgga	0	1	3	0	-
NM_000507: 38	ccctgacccgcttcgtcatggag	0	1	3	0	-
NM_000507:433	actgatgagccttctgagaagga	0	0	4	0	-
NM_000507:434	ctgatgagccttctgagaaggat	0	1	3	0	-
NM_000507:435	tgatgagccttctgagaaggatg	1	0	4	0	-
NM_000507:436	gatgagccttctgagaaggatgc	1	0	4	0	-
NM_000507:437	atgagccttctgagaaggatgct	0	1	4	0	-
NM_000507:438	tgagccttctgagaaggatgctc	0	0	3	0	-
NM_000507:439	gagccttctgagaaggatgctct	1	1	3	0	-
NM_000507:440	agccttctgagaaggatgctctg	0	1	3	0	-
NM_000507:441	gccttctgagaaggatgctctgc	1	1	4	4	1708
NM_000507:442	ccttctgagaaggatgctctgca	0	0	3	0	-
NM_000507:544	gtcaactgcttcatgctggaccc	0	1	2	0	-
NM_000507:545	tcaactgcttcatgctggacccg	0	0	2	0	-

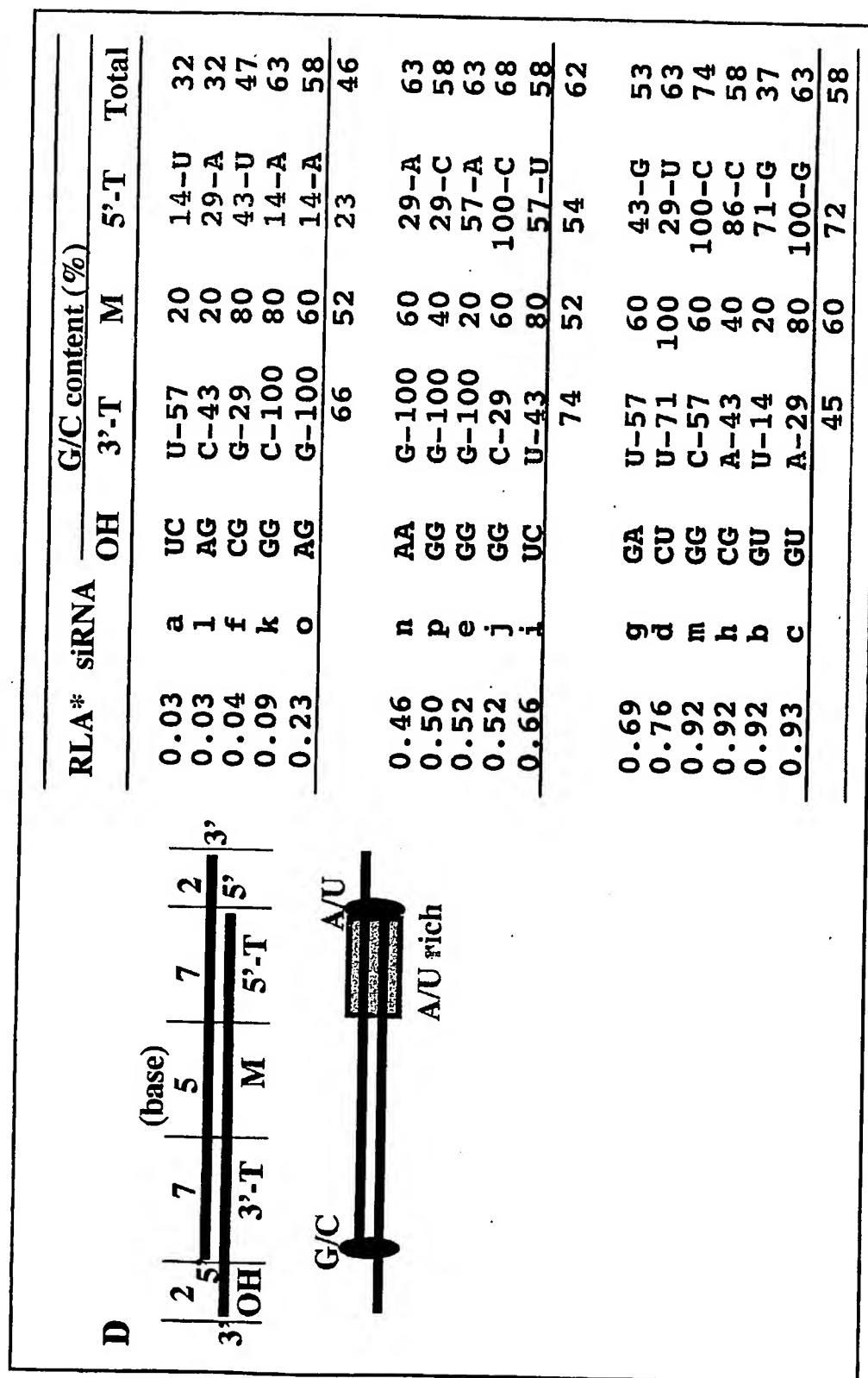
【図9】



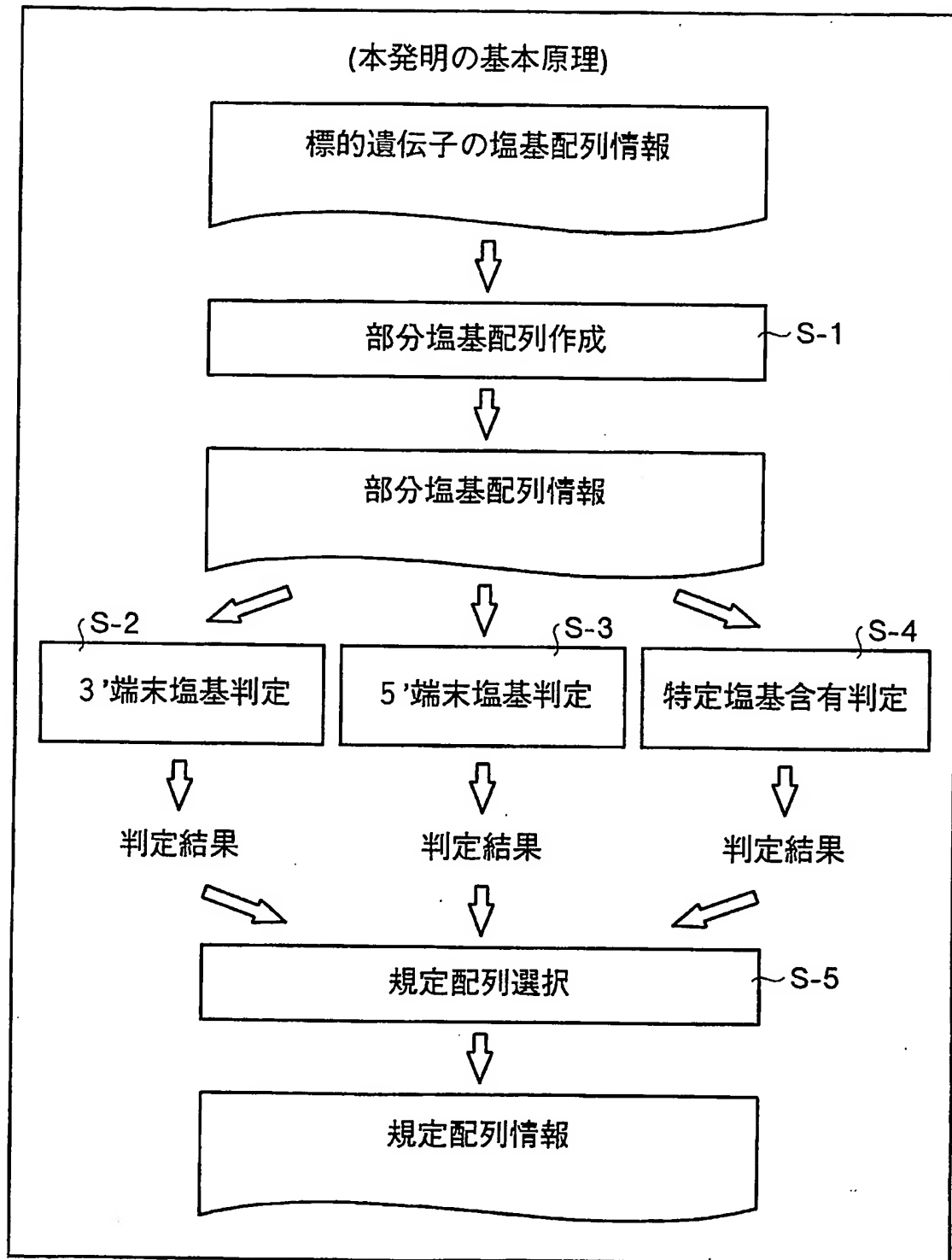
【図 10】



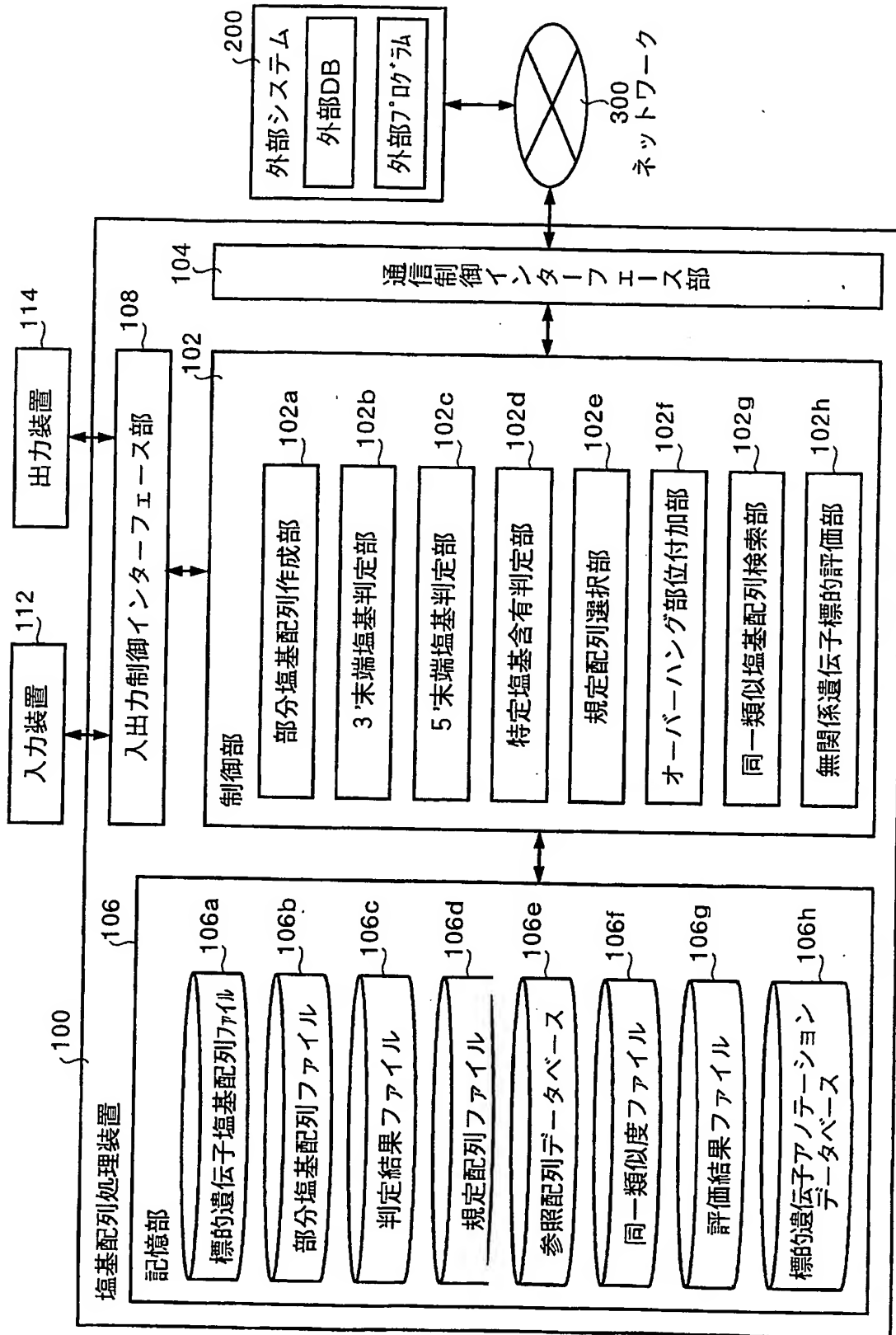
【図 11】



【図 12】



【図13】



【図14】

標的遺伝子塩基配列ファイル 106a	
塩基配列識別情報	塩基配列情報
NM_000507	ATGGCTGA ... AGTGA
⋮	⋮

【図15】

部分塩基配列ファイル 106b		
部分塩基配列 識別情報	部分塩基配列情報	オーバーハング部位 含有情報
NM_000507:36	caccct ... tcatgg	含有
⋮	⋮	⋮

【図16】

判定結果ファイル 106c				
部分塩基配列 識別情報	3'末端塩基判定 結果	5'末端塩基判定 結果	特定塩基含有判定 結果	総合判定結果
NM_000507:36	1	1	4	4
⋮	⋮	⋮	⋮	⋮



【図 17】

規定配列ファイル 106d	
部分塩基配列識別情報	規定配列情報
NM_000507:36	caccct ... tcatgg
⋮	⋮

【図 18】

参照配列データベース 106e	
参照配列識別情報	参照塩基配列情報
ref NM_015820.1	caccct ... gcatgg
⋮	⋮

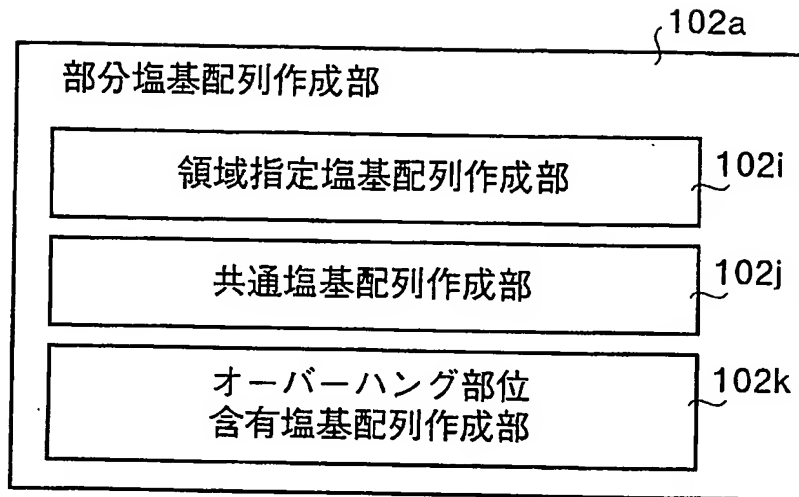
【図 19】

同一類似度ファイル 106f		
部分塩基配列 識別情報	規定識別情報	同一類似度
NM_000507:36	ref NM_015820.1	0.52
	ref NM_003837.1	0.52
⋮	⋮	⋮

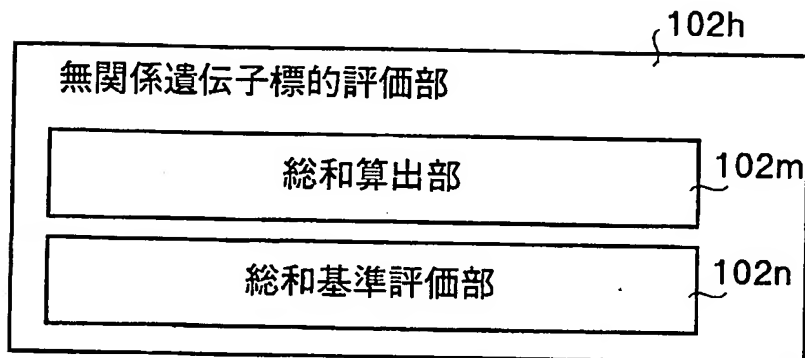
【図 20】

評価結果ファイル 106g		
部分塩基配列 識別情報	参照配列識別情報	評価情報
NM_000507:36	5.9	非標的
NM_000507:441	170.8	標的
⋮	⋮	⋮

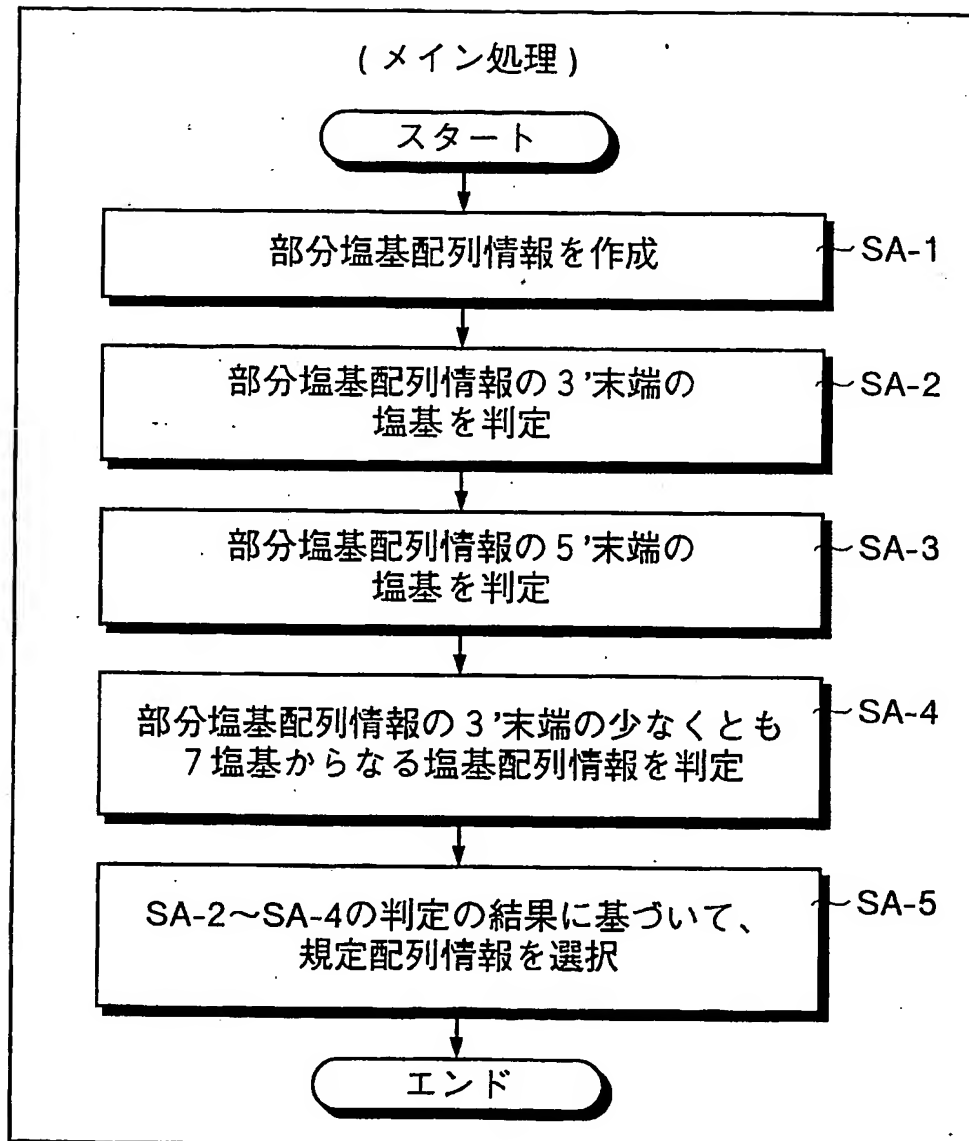
【図 21】



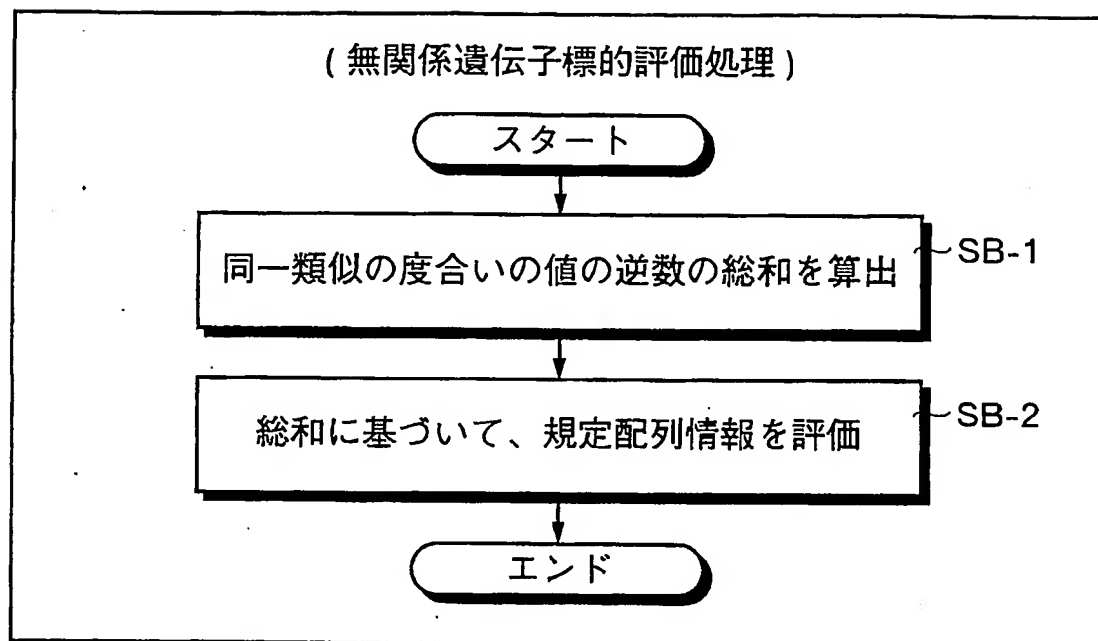
【図 22】



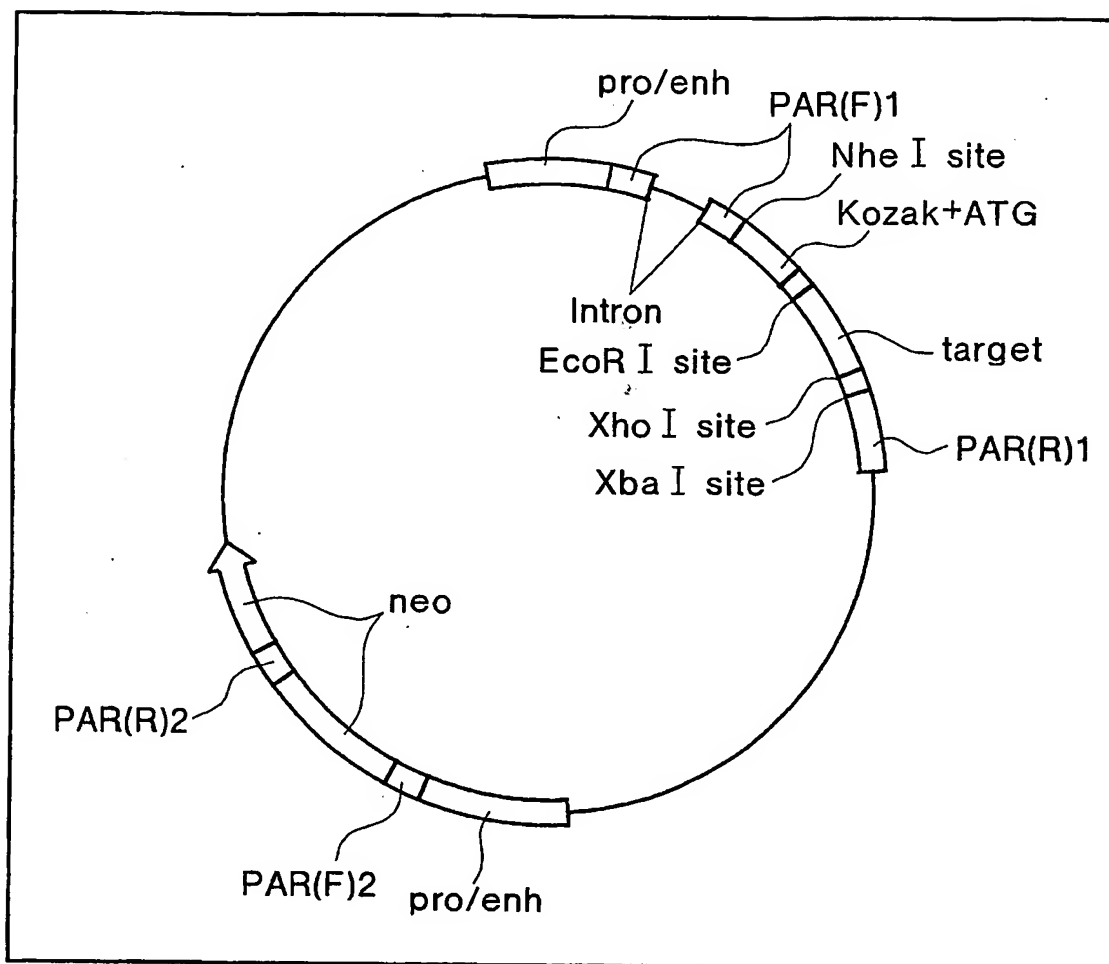
【図23】



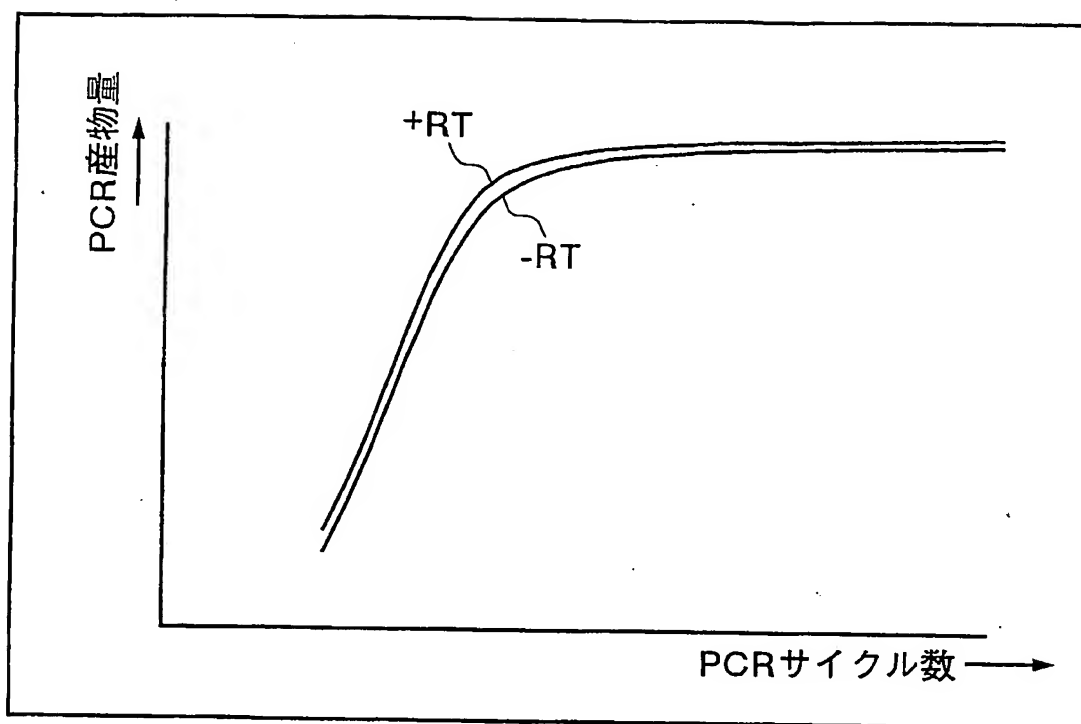
【図 24】



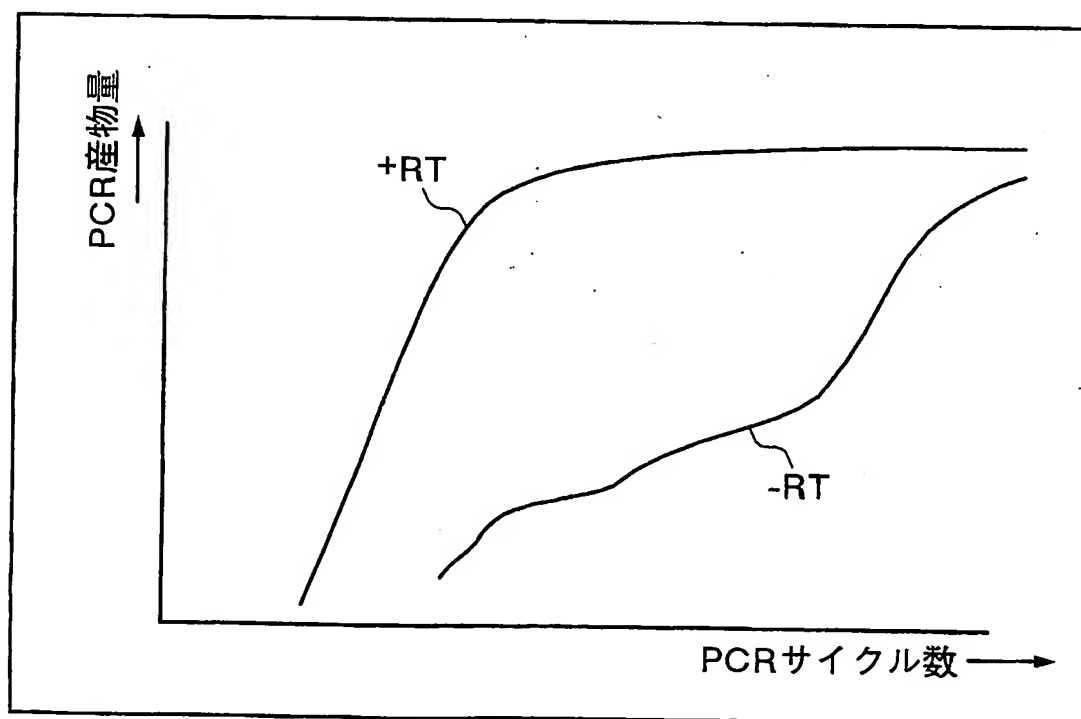
【図 25】



【図26】



【図27】



【図 2 8】

5'	-	aggauguucggcgcccgggc	-3'
3'	-	cguccuacaagccgccgggcc	-5'

【図 2 9】

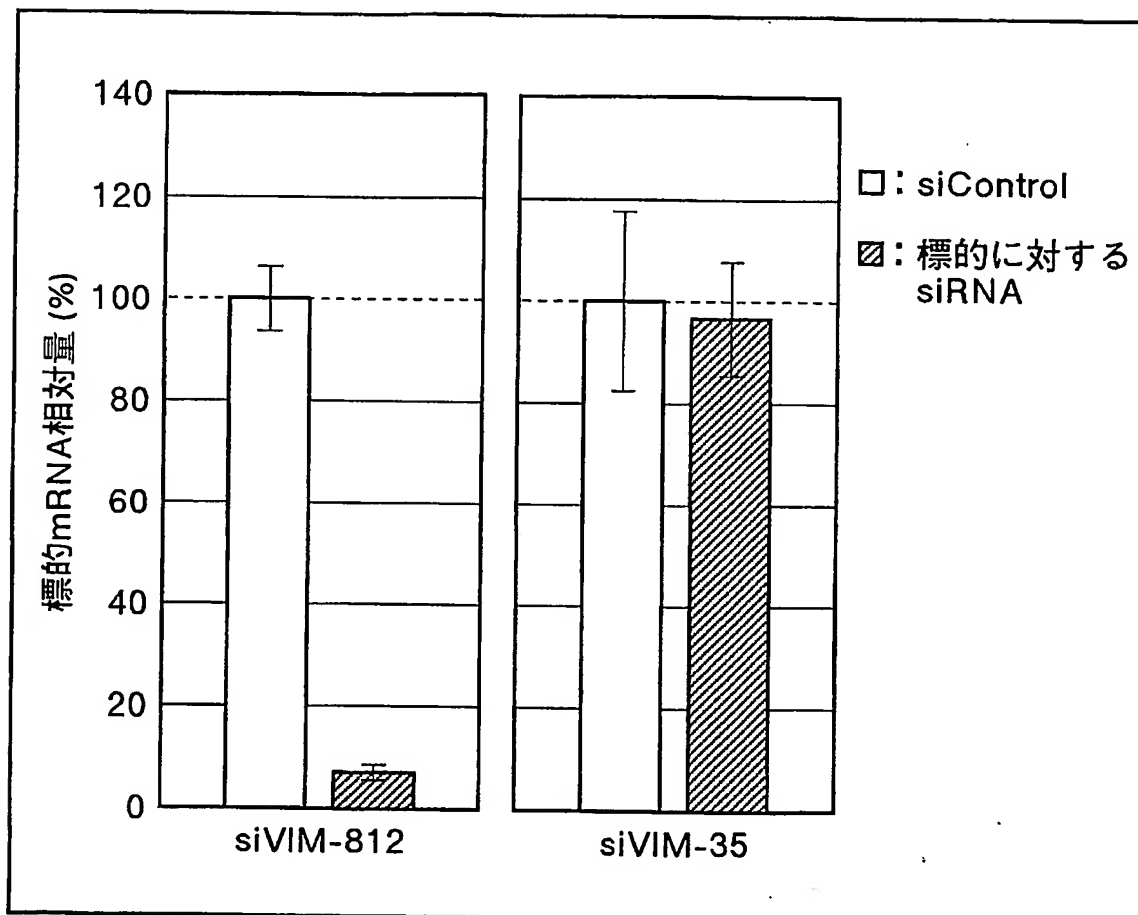
5'	-	guacgucagcaauaugaaagu	-3'
3'	-	ugcaugcagucguuauacuuu	-5'

【図 3 0】

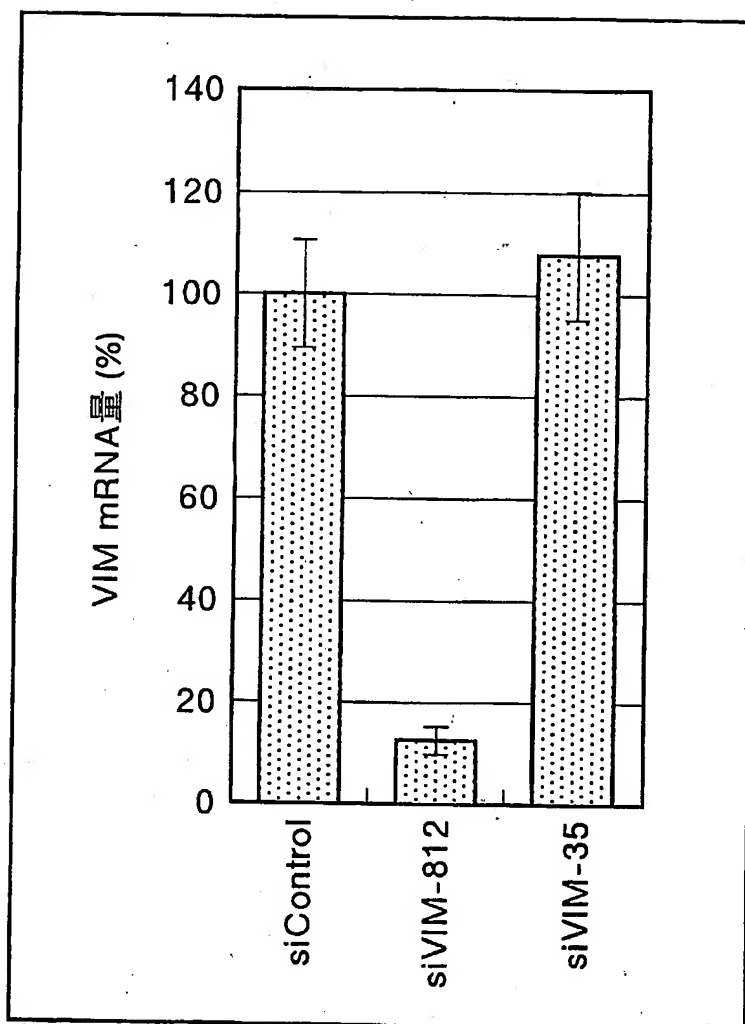
5'	-	CAUUCUAUCCGCGGAAGAUG	-3'
3'	-	CGGUAAGAUAGGCGACCUUCU	-5'



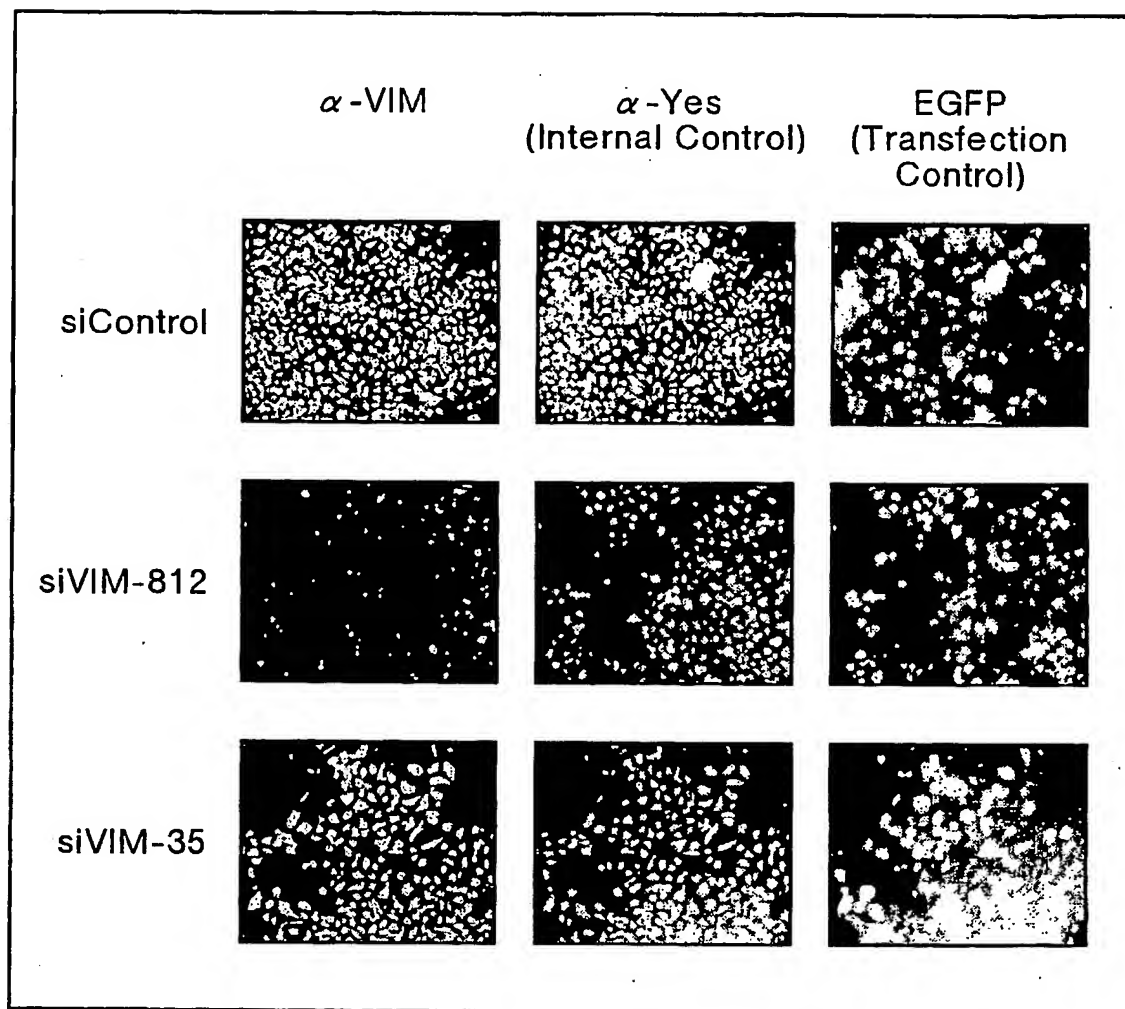
【図 31】



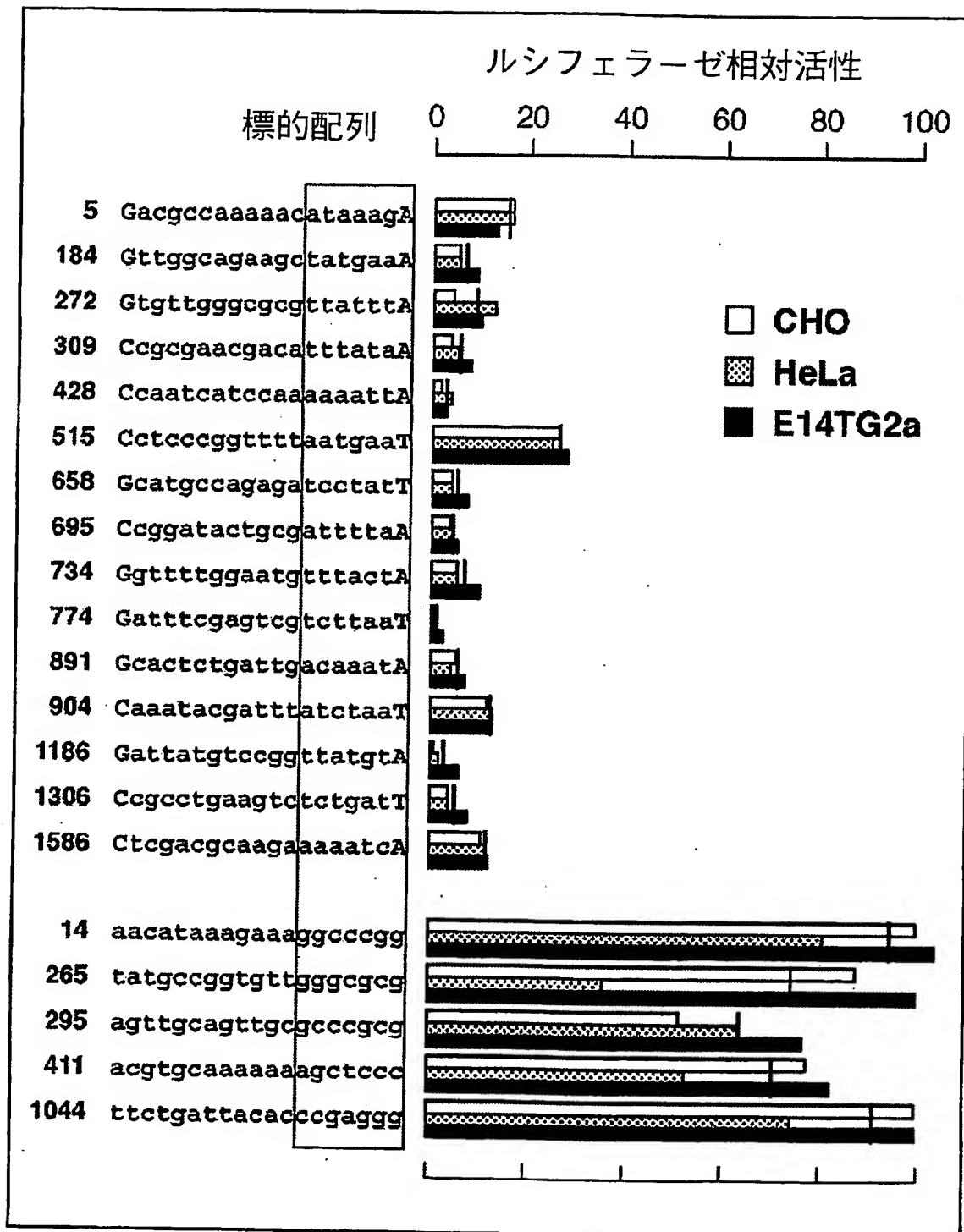
【図 3 2】



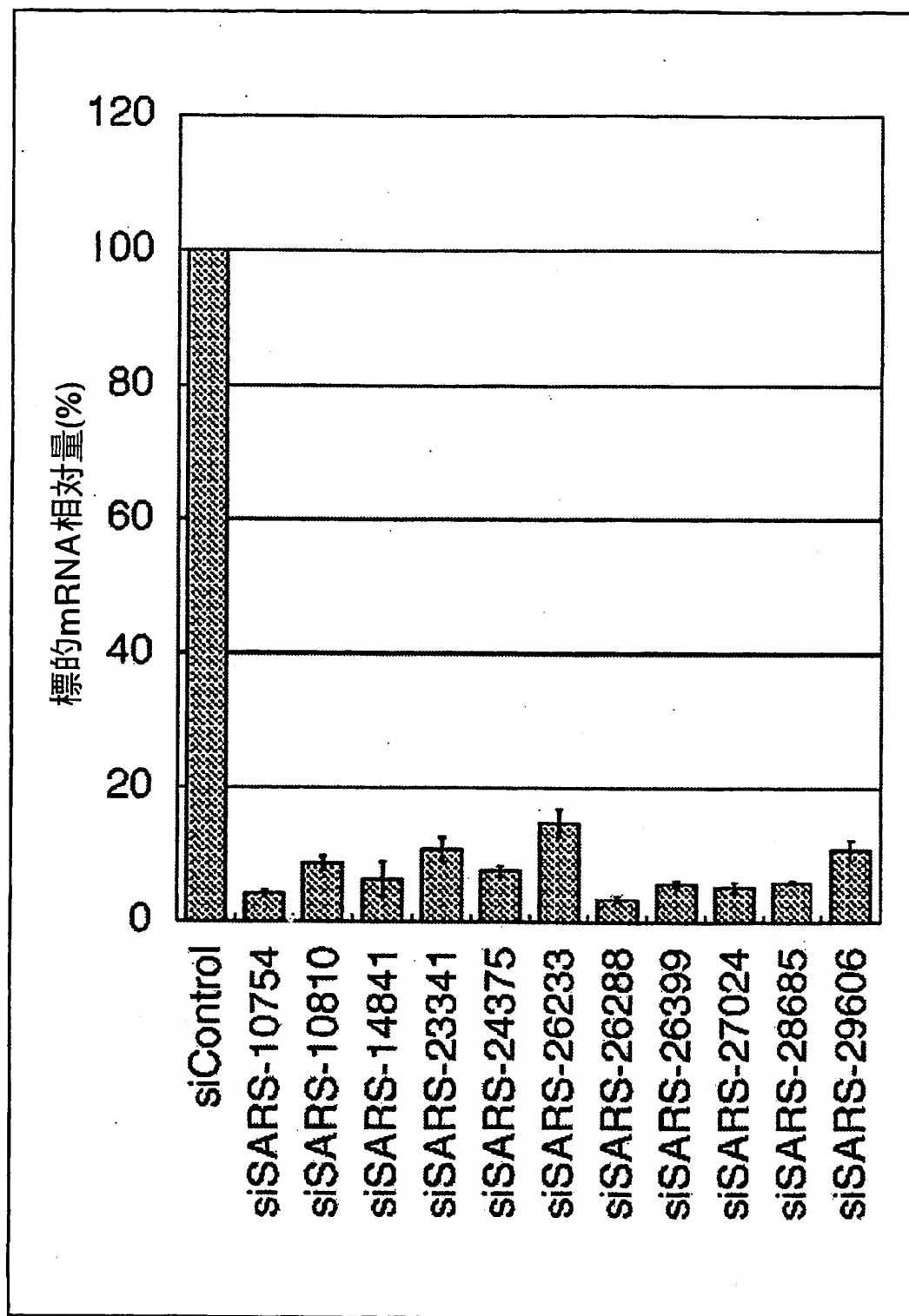
【図33】



【図34】



【図 35】





## 【書類名】要約書

## 【要約】

【課題】 標的遺伝子に対して高いRNA干渉効果を有するとともに、標的遺伝子と関係のない遺伝子についてRNA干渉を生じる危険性が非常に小さいポリヌクレオチドを提供する。

【解決手段】 RNA干渉の標的遺伝子の塩基配列から、下記(a)から(d)の規則に従う配列部位を検索し、検索結果に基づきRNAiを生じさせるポリヌクレオチドの設計、合成等を行う。

- (a) 3'末端の塩基が、アデニン、チミンまたはウラシルである。
- (b) 5'末端の塩基が、グアニンまたはシトシンである。
- (c) 3'末端の7塩基の配列において、アデニン、チミンおよびウラシルからなる群より選ばれる1種または2種以上の塩基がリッチである。
- (d) 塩基数が、細胞毒性を生じさせずにRNA干渉を生じさせ得る数である。

【選択図】 なし

特願 2004-232811

ページ: 1/E

出願人履歴情報

識別番号

[504319367]

1. 変更年月日

2004年 5月11日

[変更理由]

新規登録

住所

東京都文京区本郷三丁目32番地6-703号

氏名

株式会社RNAi

2. 変更年月日

2004年11月25日

[変更理由]

住所変更

住所

東京都千代田区内神田1-2-1 MTKビル9階

氏名

株式会社RNAi